2.
Comunicações Quânticas

Para muitos, a física quântica (também conhecida como "mecânica quântica"[4]) ainda não passa de um monte de idéias malucas e contra-intuitivas, que não poderiam e nem deveriam ser consideradas como uma descrição última da realidade. No melhor dos casos, o assunto sempre foi considerado extremamente complexo pela opinião pública, como se apenas mentes ilustres fossem capazes de compreendê-lo. Mesmo tendo sido responsável por uma revolução sem precedentes na indústria eletro-eletrônica, devido ao advento dos dispositivos semicondutores, esse paradigma não se alterou muito com o passar dos tempos, o que pode soar estranho à primeira vista. Como uma teoria tão bem sucedida pode ser, ainda nos dias de hoje, encarada como bruxaria?

A resposta mais provável é que, apesar da física quântica ter surgido no início do século XX, a humanidade ainda não viu nenhuma aplicação prática de suas teorias mais radicais, isto é, que explicam o comportamento de átomos e partículas isolados. A física de lasers e semicondutores, responsáveis pela revolução citada acima, não passa de manifestações da teoria quântica em aglomerados de átomos; isso significa que ela pode ser compreendida através de modelos semi-clássicos, que não fogem muito ao paradigma determinístico com o qual estamos familiarizados. Portanto, ainda era necessário surgir uma aplicação das teorias verdadeiramente "quânticas" para que essa discussão se aproximasse do restante da sociedade. E ela não apenas já surgiu como está progredindo a pleno vapor!

Chamamos de "Comunicações Quânticas" uma sub-área da teoria da informação quântica que lida com a transmissão, propagação e detecção dos portadores quânticos da informação, os chamados *quantum bits* ou *qubits*. O conceito do qubit é, talvez, o que há de mais revolucionário no advento da teoria

[4] Alguns pesquisadores, entre eles Nicolas Gisin, defendem que "mecânica quântica" foi um erro histórico de nomenclatura. Esse ponto de vista está baseado no fato de que, pura e simplesmente, não existe nada de "mecânico" na teoria! Desta forma, essa nomenclatura, embora largamente utilizada, não é utilizada nesta tese.

da informação quântica. Embora ele não seja nada mais que um sistema quântico de dois níveis, assim como o *bit* clássico é um sistema clássico de dois níveis, o qubit forneceu, pela primeira vez, uma interpretação *física* para a informação, e o nome passou a ser usado tanto para descrever o grau de liberdade (de 2 níveis) no qual a informação foi codificada como para também descrever o próprio portador da informação. Para comunicações de longas distâncias, a radiação eletro-magnética é a candidata mais natural para se implementar qubits, e o fóton surge como "partícula elementar" de comunicações quânticas. Esse fato dá uma posição extremamente privilegiada a todos os grupos de pesquisa do mundo que trabalham na área de óptica – ou, mais especificamente, *óptica quântica*.

Este capítulo tem como objetivo fazer uma breve introdução ao assunto, e derivar algumas expressões que são utilizadas no restante da tese. A seção 2.1 explica, de um ponto de vista matematicamente formal, o que é um qubit. Em seguida, os três componentes principais dos sistemas de telecomunicações – fonte, canal e receptor – são discutidos do ponto de vista da óptica quântica, respectivamente nas seções de 2.2 a 2.4. Finalmente, a seção 2.5 faz uma introdução à aplicação mais importante, e mais bem sucedida, da teoria da informação quântica: a criptografia quântica. É importante ressaltar que todas as demais aplicações da óptica quântica, tais como o teletransporte quântico, a memória quântica e o *entanglement swapping*, foram deixados de fora; para maiores informações, ver as referências [4,5].

2.1. Qubits

A entidade mais fundamental na teoria da informação é o bit. Qualquer sistema físico que possua dois estados distintos pode ser utilizado para codificar informação nos bits "0" e "1", como por exemplo um interruptor de luz que pode se encontrar nos estados "ligado" ou "desligado".

Na teoria da informação quântica, o análogo do bit é o chamado *quantum bit*, ou abreviadamente *qubit*. Qubits são objetos estranhos! Apesar de haver diversas similaridades, o qubit é uma entidade muito diferente do bit clássico. Para melhor compreendermos essa diferença, vamos começar essa seção pela definição formal de qubit.

Definição 1. *Um qubit é um vetor de estado em um espaço de estados bi-dimensional, ou seja, um vetor unitário* $|\Psi\rangle \in \mathcal{H}$ *onde* $\mathcal{H}$ *é um espaço de Hilbert complexo de dimensão 2.*

Em outras palavras, um qubit é um estado qualquer de um sistema quântico de duas dimensões[5]. Embora essa definição seja suficiente para caracterizar um qubit, ela fornece pouca intuição. Talvez a idéia fique mais clara quando conseguirmos fazer alguma analogia com os "zeros" e "uns" dos bits, o que é feito na próxima sub-seção.

2.1.1. Representação de qubits

Existem diferentes formas de se representar qubits. Uma das mais usuais consiste em escrevê-los como combinações lineares de uma certa base ortonormal, chamada de *base computacional*, representada pelos kets $|0\rangle$ e $|1\rangle$[6]. Esta base está normalmente associada a um dos observáveis que representam o conjunto de medidas que podem ser realizadas nos qubits, e atribuímos a cada componente da base os valores (clássicos) 0 e 1, isto é,

$$\begin{aligned} 0 &\mapsto |0\rangle \\ 1 &\mapsto |1\rangle \end{aligned} \tag{2.1}$$

Apenas para dar um exemplo, um qubit poderia ser implementado por um sistema atômico de 2 níveis, no qual $|0\rangle$ corresponderia, por exemplo, ao estado fundamental do átomo, e $|1\rangle$ ao estado excitado. Mas, alguém poderia perguntar, qual a diferença entre um sistema quântico desse tipo e uma lâmpada que pode estar acesa ou apagada? A resposta é que sistemas quânticos podem se encontrar

[5] Ver o Apêndice A para uma breve introdução a espaços de Hilbert.

[6] Atenção: não confundir o ket $|0\rangle$ com o vetor nulo! Além disso, em algumas partes dessa tese, os kets $|0\rangle$ e $|1\rangle$ serão utilizados para representar os estados de Fock com 0 e 1 fóton.

em *superposições* de estados, ou seja, em *combinações lineares* entre os estados $|0\rangle$ e $|1\rangle$. Desta forma, um qubit genérico pode ser escrito como:

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle \tag{2.2}$$

onde α e β são números complexos que representam as amplitudes de probabilidade de se medir cada um dos valores[7], isto é,

$$p("0") = |\langle 0|\Psi\rangle|^2 = |\alpha|^2, \; p("1") = |\langle 1|\Psi\rangle|^2 = |\beta|^2 \tag{2.3}$$

onde $|\alpha|^2 + |\beta|^2 = 1$. A notação (2.2) sugere que qualquer qubit é dado por uma superposição coerente de dois estados quânticos – ou seja, um qubit pode assumir quaisquer valores intermediários entre "0" e "1" em um espaço contínuo parametrizado pelos coeficientes α e β. Por exemplo, se um qubit for dado por

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \tag{2.4}$$

e ele for medido na base computacional, existe uma chance de 50% de ser medido como "0" e 50% de ser medido como "1". Ou seja, ele é, de certa forma, "0" e "1" ao mesmo tempo! No exemplo do átomo de 2 níveis, é como se o átomo se encontrasse ao mesmo tempo excitado e relaxado, o que nos parece estranho mas é perfeitamente válido pelas leis da física quântica.

No entanto, também de uma certa forma, qualquer qubit sempre possui um valor definido, dependendo de em qual base realizamos a medida. Seja, por exemplo, a base obtida pela aplicação de uma transformação *H* à base computacional:

[7] Se a medida for realizada na base computacional.

$$0 \mapsto |+\rangle = H|0\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)$$
$$1 \mapsto |-\rangle = H|1\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right) \qquad (2.5)$$

onde H é chamada de *transformação de Hadamard*. De acordo com (2.5), o ket $|+\rangle$ representa o valor lógico "0" e $|-\rangle$ o valor lógico "1". É fácil ver que, nessa base, o estado de (2.4) tem valor definido e igual a 0. É interessante notar que, na prática, nem sempre é simples, de um ponto de vista experimental, realizar uma mudança de base como a de (2.5), vide o exemplo do átomo de 2 níveis.

Para facilitar a visualização do espaço em que se encontra um qubit, escrevemos, sem perda de generalidade, $\alpha = e^{j\gamma}\cos(\theta/2)$ e $\beta = e^{j\delta}\sin(\theta/2)$. Substituindo em (2.2), obtemos:

$$|\Psi\rangle = e^{j\gamma}\left(\cos\frac{\theta}{2}|0\rangle + e^{j\phi}\sin\frac{\theta}{2}|1\rangle\right) \qquad (2.6)$$

onde $\phi = \delta - \gamma$. Dado que a fase global γ não produz efeitos observáveis, podemos ignorá-la. Os números θ, ϕ parametrizam a superfície de uma esfera, chamada *esfera de Bloch*[8]. Nessa representação, os pólos da esfera correspondem aos estados $|0\rangle$ e $|1\rangle$, como mostra a figura 1.

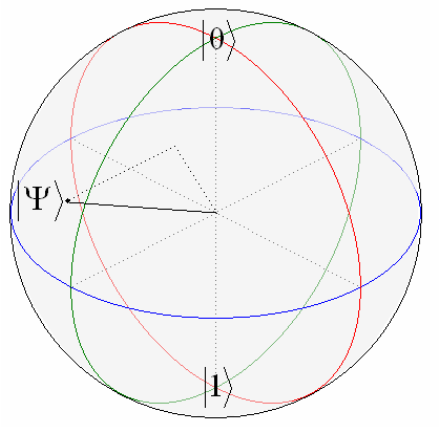

Figura 1: Representação de qubits como pontos na superfície de uma esfera.

[8] Também conhecida como *esfera de Poincaré* devido à direta associação que podemos fazer entre um vetor de estado e o vetor de Jones para a descrição da luz polarizada.

Observe que, na representação pela esfera de Bloch, estados ortogonais encontram-se diametralmente opostos. Essa representação mostra claramente que um qubit pode estar em qualquer um dos infinitos pontos da superfície da esfera; no entanto, embora a quantidade de informação clássica necessária para se descrever um qubit seja infinita, só podemos extrair 1 bit clássico de informação. O motivo para isso está no simples fato de que não podemos obter nenhuma informação sem realizar uma medida, e quando a medida é feita, sempre obtemos como resultado um dos auto-estados $|0\rangle$ ou $|1\rangle$. Ou seja, para nosso desgosto, a "infinita informação" contida em um qubit é inacessível!

2.1.2. Múltiplos qubits

Suponha agora que desejamos representar um conjunto de *n* qubits, da mesma forma que podemos escrever seqüências de bits clássicos, como por exemplo "01" e "110". Dado que cada qubit é representado por um sistema físico distinto (átomos, fótons, etc), o estado global é obtido pelo produto tensorial de cada estado individual, ou seja,

$$|\Psi\rangle = |\Psi_1\rangle \otimes |\Psi_2\rangle \otimes \cdots \otimes |\Psi_n\rangle \tag{2.7}$$

onde $|\Psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_n$. Em alguns casos, para a notação não ficar desnecessariamente carregada, escrevemos simplesmente $|\Psi\rangle = |\Psi_1\rangle|\Psi_2\rangle \cdots |\Psi_n\rangle$, onde o produto tensorial fica subentendido.

Uma propriedade muito interessante de múltiplos qubits está no fato de que, assim como um qubit simples pode se encontrar em uma superposição dos dois estados mensuráveis $|0\rangle$ e $|1\rangle$, um conjunto de qubits também pode se encontrar em uma superposição de todas as possíveis combinações desses estados. Por exemplo, se $|\Psi\rangle = |0\rangle|0\rangle$, a aplicação da transformação de Hadamard (2.5) a cada qubit resulta em

$$H|0\rangle \otimes H|0\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \otimes \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)$$
$$= \frac{1}{2}\left(|0\rangle|0\rangle + |0\rangle|1\rangle + |1\rangle|0\rangle + |1\rangle|1\rangle\right) \tag{2.8}$$

que é uma superposição de todos os 4 possíveis valores clássicos "0" e "1". Esse resultado pode ser naturalmente estendido para mais de 2 qubits: a aplicação de uma mesma transformação em *n* qubits pode resultar em uma superposição de todos os 2^n possíveis valores. Podemos interpretar esse fato como uma primeira vantagem de um computador quântico sobre um computador clássico, que é normalmente chamada de *paralelismo quântico*: se o registrador se encontrar em uma superposição de todos os valores possíveis, uma tarefa (digamos, calcular os valores de uma função para todas as entradas possíveis) pode ser realizada em um tempo exponencialmente mais curto que em um computador clássico![9]

2.1.3. Emaranhamento

O princípio da superposição prevê a possibilidade de dois ou mais sistemas quânticos se encontrarem em um estado que não pode ser escrito da forma (2.7). Chamamos esses estados não-separáveis de *emaranhados*[10].

Para facilitar a ilustração, considere um sistema de 2 qubits *A* e *B*, cujos espaços de Hilbert $\mathcal{H}_A$ e $\mathcal{H}_B$ têm os vetores $\left\{|0\rangle_A, |1\rangle_A\right\}$ e $\left\{|0\rangle_B, |1\rangle_B\right\}$ como base. O estado mais genérico pertencente a $\mathcal{H}_A \otimes \mathcal{H}_B$ é dado por

$$|\Psi\rangle_{AB} = \sum_{i,j} c_{ij} |i\rangle_A |j\rangle_B \tag{2.9}$$

---

[9] É claro que, assim como no caso de um qubit simples, não temos acesso direto à informação contida no estado final do registrador quântico, sendo necessários algoritmos inteligentes o suficiente para tirar vantagem desse processo. Os algoritmos de Shor e Grover são bons exemplos.

[10] Na literatura disponível em português, o termo universal *entanglement* foi também traduzido, por alguns autores, como "entrelaçamento".

onde $i, j \in \{0,1\}$. Estados emaranhados são estados do tipo (2.9) que não podem ser escritos como um produto tensorial entre um estado em *A* e outro em *B*, ou seja, da forma $\sum_i c_i |i\rangle_A \otimes \sum_j c_j |j\rangle_B$.

Considere o caso, por exemplo, do seguinte estado:

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B\right) \tag{2.10}$$

É fácil observar que este estado é emaranhado; de fato, por razões que transcendem o escopo dessa tese, ele é um tipo de estado *maximamente emaranhado*.

Estados emaranhados como o descrito por (2.10) são caracterizados por exibir uma *correlação* entre os sistemas componentes. Observe que, se o qubit *A* for medido e o resultado da medida for "0", isso implica que o resultado de uma medida no qubit *B* deve, necessariamente, resultar no valor "1", e vice-versa.

O emaranhamento é uma característica unicamente quântica, sem nenhum análogo na física clássica, e responsável por uma discussão quase interminável que surgiu em 1935 com um paper de Einstein, Podolsky e Rosen que, partindo de um experimento imaginário envolvendo elétrons emaranhados, tentava demonstrar que a física quântica é uma "teoria incompleta da realidade" [6]. A conclusão de EPR, no entanto, assume a hipótese de *realismo local*, que, desde a publicação de um artigo de John Bell em 1968, sabemos ser incompatível com as previsões da física quântica [7].

2.1.4. Estados mistos e a matriz de densidade

Embora a equação (2.2) seja muito útil para representar qubits cujos estados são perfeitamente conhecidos (chamados de estados *puros*), ela não permite uma descrição de situações nas quais o estado do sistema é dado por uma superposição incoerente de vários estados, os chamados estados *mistos*[11].

[11] Na analogia com a descrição da polarização da luz, estados "puros" correspondem à luz polarizada e estados "mistos" à luz parcialmente polarizada.

Para ilustrar o problema, suponha que um certo qubit, cujo estado $|\Psi\rangle$ desconhecemos, pode se encontrar em uma combinação qualquer de um certo número de estados $|\psi_i\rangle$ com respectivas probabilidades p_i. Para representar essa situação, utilizamos a chamada *matriz de densidade*, definida como:

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \tag{2.11}$$

Observe que a matriz de densidade de um estado puro $|\Psi\rangle$ é simplesmente dada por $|\Psi\rangle\langle\Psi|$, ou seja, a equação (2.11) é mais genérica que a representação (2.2), exclusiva de estados puros. Mas o que ela traz de vantagem com relação à (muito mais simples) representação anterior? Para respondermos a essa pergunta, considere, por exemplo, a matriz de densidade dada por:

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}I \tag{2.12}$$

A expressão (2.12) representa uma situação na qual os estados $|0\rangle$ e $|1\rangle$ encontram-se em uma superposição sem qualquer relação de fase. É evidente de (2.12) que, se uma medida na base computacional for realizada, há probabilidades de 50% de se obter cada um dos resultados "0" e "1". Uma pergunta que surge freqüentemente é: o que diferencia o estado (2.12) do estado $|+\rangle$ de (2.5)? A resposta está no fato que (2.5) é uma superposição *coerente* dos estados descritos pela base computacional, enquanto (2.12) representa uma superposição *incoerente*. Isso pode ser facilmente percebido tomando como exemplo os estados (2.5); nesse caso, a matriz de densidade seria dada por:

$$\begin{aligned}\rho &= \frac{1}{2}\left(|+\rangle\langle +| + |-\rangle\langle -|\right) \\ &= \frac{1}{2}\left[\frac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1| + |1\rangle\langle 0| + |0\rangle\langle 1|\right) + \frac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1| - |1\rangle\langle 0| - |0\rangle\langle 1|\right)\right] \\ &= \frac{1}{2}|0\rangle\langle 0| + |1\rangle\langle 1|\end{aligned} \tag{2.13}$$

Isso confirma o fato de (2.12) representar uma mistura incoerente. Essa mistura é muitas vezes chamada de mistura de *máxima incerteza*, pois, antes de se realizar uma medida, não temos nenhuma informação *a priori* sobre o estado no qual o qubit se encontra.

Uma outra forma de se chegar ao resultado (2.13) é aplicarmos a transformação de Hadamard diretamente em (2.12). Mas como aplicamos transformações unitárias em matrizes de densidade? Basta usar a definição (2.11) e observar que:

$$\rho' = \sum_i p_i U|\psi_i\rangle\langle\psi_i|U^* = U\left(\sum_i p_i|\psi_i\rangle\langle\psi_i|\right)U^* = U\rho U^* \tag{2.14}$$

É trivial observar que, no caso de ρ ser dada por (2.12), nenhuma transformação unitária é capaz de modificar o estado do sistema. Esse fato demonstra que a mesma matriz de densidade (2.10) pode ser obtida considerando-se uma mistura incoerente de *quaisquer* estados ortogonais na proporção 50%/50%.

Existem algumas propriedades da matriz de densidade que podem ser úteis na resolução de problemas, que são listadas a seguir[12].

i. (*Unidade do traço*) ρ possui traço igual a 1.

ii. (*Positividade*) ρ é um operador positivo, isto é, com autovalores λ_j tais que $\lambda_j \geq 0 \ \ \forall j$.

iii. (*Critério dos estados puros*) $\mathrm{tr}\left(\rho^2\right) \leq 1$, com igualdade se, e somente se, ρ descrever um estado puro.

Matrizes de densidade também são muito úteis para a descrição de múltiplos qubits. Seguindo a lógica da expressão (2.7), a matriz de densidade de um sistema contendo *n* qubits é dado pelo produto tensorial das matrizes de densidade de cada sub-sistema, isto é,

[12] Para a demonstração de todas as propriedades, ver a referência [8].

$$\rho = \rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n \tag{2.15}$$

onde pressupõe-se, evidentemente, que (2.15) é um estado produto. Mas e no caso de haver emaranhamento? Dessa simples pergunta surge o que talvez seja a aplicação mais importante das matrizes de densidade: a descrição de *sub-sistemas* de um sistema quântico composto.

Suponha que dois sistemas físicos *A* e *B* são descritos por uma matriz de densidade ρ^{AB}. A *matriz de densidade reduzida* para o sistema *A* é definida da forma:

$$\rho^A = \mathrm{tr}_B\left(\rho^{AB}\right) \tag{2.16}$$

onde tr_B é chamado de *traço parcial sobre o sistema B* e é definido como:

$$\mathrm{tr}_B\left(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|\right) = |a_1\rangle\langle a_2|\,\mathrm{tr}\left(|b_1\rangle\langle b_2|\right) \tag{2.17}$$

onde $|a_i\rangle \in \mathcal{H}_A, |b_i\rangle \in \mathcal{H}_B$. É possível mostrar que o traço parcial deve ser utilizado *sempre* que se desejar obter a matriz de densidade de um sub-sistema. Esse fato não é nem um pouco evidente, mas pode ser facilmente verificado para o caso de estados produto. Considerando $\rho^{AB} = \rho^A \otimes \rho^B$, temos

$$\mathrm{tr}_B\left(\rho^{AB}\right) = \mathrm{tr}_B\left(\rho^A \otimes \rho^B\right) = \rho^A\,\mathrm{tr}\left(\rho^B\right) = \rho^A \tag{2.18}$$

que é o resultado esperado. Um exemplo um pouco mais complexo, mas muito mais esclarecedor, é o caso dos estados emaranhados. No caso do estado $|\Psi^+\rangle$ da equação (2.10), temos que:

$$\begin{aligned}\rho^{AB} &= \left(\frac{|01\rangle + |10\rangle}{\sqrt{2}}\right)\left(\frac{\langle 01| + \langle 10|}{\sqrt{2}}\right) \\ &= \frac{1}{2}\left(|01\rangle\langle 01| + |10\rangle\langle 10| + |01\rangle\langle 10| + |10\rangle\langle 01|\right)\end{aligned} \tag{2.19}$$

Se desejarmos obter uma descrição do estado do qubit A, tomamos o traço parcial com relação a B:

$$\begin{aligned}\rho^A &= \mathrm{tr}_B\left(\rho^{AB}\right) \\ &= \frac{1}{2}\left(\mathrm{tr}_B|01\rangle\langle 01| + \mathrm{tr}_B|10\rangle\langle 10| + \mathrm{tr}_B|01\rangle\langle 10| + \mathrm{tr}_B|10\rangle\langle 01|\right)\end{aligned} \tag{2.20}$$

Usando (2.17) juntamente ao fato que $\mathrm{tr}\left(|b_1\rangle\langle b_2|\right) = \langle b_2|b_1\rangle$, temos:

$$\begin{aligned}\rho^A &= \frac{1}{2}\left(|0\rangle\langle 0|\langle 1|1\rangle + |1\rangle\langle 1|\langle 0|0\rangle + |0\rangle\langle 1|\langle 0|1\rangle + |1\rangle\langle 0|\langle 1|0\rangle\right) \\ &= \frac{1}{2}\left(|0\rangle\langle 0| + |1\rangle\langle 1|\right) = \frac{1}{2}I\end{aligned} \tag{2.21}$$

Esse resultado mostra uma importantíssima característica dos estados emaranhados: enquanto o estado global é um estado *puro*, a observação de apenas um dos sub-sistemas resulta em um estado *misto*. Além disso, um raciocínio reverso sobre todas as etapas (2.19) a (2.21) mostra que é possível representar estados mistos como sub-sistemas de um sistema de dimensão maior no qual o estado é puro, processo conhecido pelo nome de *purificação*.

2.1.5. Implementações práticas

Já vimos que qualquer sistema quântico de dois estados é, em princípio, capaz de implementar qubits. Alguns, porém, são mais práticos e eficientes nessa tarefa. No caso de comunicações quânticas, em que estamos interessados na transmissão de informação através de longas distâncias, o candidato perfeito é a radiação eletromagnética, isto é, o *fóton*.

Mas por acaso um fóton é um sistema de dois níveis? Em geral, não! Por isso, precisamos escolher apenas um *grau de liberdade* do fóton que se comporte como um sistema de dois níveis para codificar a informação quântica. As formas mais usuais são a *polarização* e a *fase* dos fótons presentes em um pulso, embora

a codificação em freqüência também já tenha sido implementada experimentalmente [9]. Recentemente, no contexto de criptografia quântica, foi sugerida uma codificação em pares de pulsos consecutivos contendo, cada um, estados coerentes com 0 ou μ fótons ($\mu < 1$) e que possuem uma relação de fase entre si [10]. Todas essas codificações podem ser facilmente implementadas em fótons gerados por lasers atenuados (ver seção 2.2); no caso de sistemas de comunicações quânticas usando pares de fótons emaranhados, que estão fora do escopo dessa tese, também é possível utilizar o par posição-momento [11].

2.1.5.1. Codificação em polarização

A escolha da polarização para implementar qubits é totalmente natural, já que a polarização do fóton é, "de fábrica", um sistema quântico de 2 níveis. De fato, é impossível não fazer uma analogia direta entre a esfera de Bloch da figura 1 e a esfera de Poincaré usada para descrever estados de polarização da luz, já que os vetores de Jones possuem exatamente a forma (2.6)[13]. Podemos, por exemplo, identificar a base computacional $\left(|0\rangle,|1\rangle\right)$ com os estados de polarização linear vertical e horizontal $\left(|\updownarrow\rangle,|\leftrightarrow\rangle\right)$.

Uma grande vantagem de se utilizar a polarização para codificar qubits está na enorme facilidade de se gerar um estado de polarização genérico. Dado que o feixe de um laser é naturalmente polarizado, é preciso apenas fazer com que o fóton passe através de componentes de birrefringência variável para que sua polarização seja variada. Além disso, caso deseje-se medir um qubit em uma base diferente da base computacional – por exemplo, na base de Hadamard (2.5) que corresponde aos estados de polarização lineares +45º e -45º – bastaria usar um componente semelhante na detecção.

A figura 2 ilustra esquematicamente como qubits codificados em polarização podem ser preparados e medidos. Supõe-se que o fóton na entrada, à esquerda da figura, se encontre em um estado de polarização conhecido, como por exemplo o estado $|0\rangle$.

[13] Para maiores detalhes sobre a representação de estados de polarização, ver o capítulo 2 de [12] ou, para um tratado completo, [13].

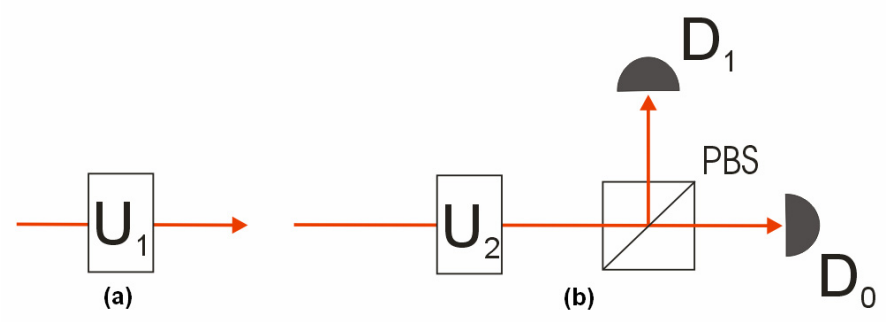

Figura 2. Esquemas para (a) Preparação e (b) Medida de qubits codificados em polarização. Ver texto para detalhes.

Para efetuar transformações no estado de polarização de um fóton, é preciso fazê-lo atravessar um elemento birrefringente, que está representado pelos componentes U_i nas figuras 2(a) e 2(b) (a notação vem do fato que trata-se de uma transformação unitária). Esse elemento pode ser implementado de diversas formas. A mais simples, porém menos eficiente, como veremos a seguir no contexto da criptografia quântica, seria utilizar uma lâmina de meia-onda seguida de uma lâmina de quarto de onda, com ambas sendo capazes de girar em torno de um eixo central. No entanto, é mais comum utilizar-se componentes eletro-ópticos, como por exemplo células de Pockels ou atuadores piezoelétricos. É possível mostrar[14] que todos esses esquemas são capazes de reproduzir todos os estados de polarização.

No estágio de detecção (medida), observe a presença de um separador de feixes polarizador (PBS, *polarizing beamsplitter*). Sua função é extremamente simples: se o fóton em sua entrada possuir polarização vertical, ele é transmitido e detectado por D_0; no entanto, se sua polarização for horizontal, ele é refletido e detectado por D_1. Obviamente, se o fóton na entrada do PBS estiver em uma superposição de estados de polarização, ele poderá ir para qualquer uma das duas direções, com probabilidades que dependerão de seu estado de polarização, ou seja, das amplitudes de probabilidade α, β e da configuração do dispositivo birrefringente.

[14] Para as demonstrações de que, de fato, o elemento birrefringente é capaz de gerar um estado de polarização arbitrário, ver [12].

2.1.5.2.
Codificação em “time bins”

A idéia de codificar qubits na fase de um fóton foi mencionada pela primeira vez por Bennett no artigo em que ele apresentou seu protocolo de criptografia quântica de dois estados (hoje chamado de B92) [14]. A codificação em “time bins” utiliza uma mescla de codificação em fase e tempo de chegada, que foi proposta em [15]. Nessa implementação, que hoje em dia também soa natural, os qubits são preparados e medidos utilizando-se *interferômetros*. A figura 3 ilustra uma situação na qual um interferômetro de Mach-Zender (implementado em fibras ópticas) é utilizado.

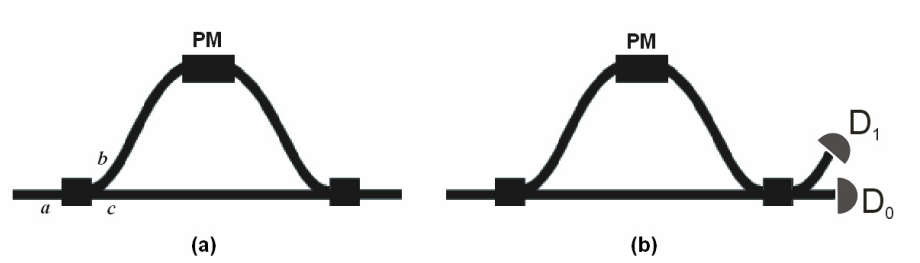

Figura 3. (a) Preparação e (b) Medida de qubits codificados em time-bins

Suponha que o fóton no qual desejamos codificar a informação entra no interferômetro pela porta *a*. Chamando esse modo de $|a\rangle$, vemos que o efeito do acoplador é dado pela transformação $|a\rangle \rightarrow c_1|b\rangle + jc_2|c\rangle$, onde os modos $|b\rangle$ e $|c\rangle$ correspondem às saídas do acoplador. Dado que o interferômetro é desbalanceado, os percursos podem ser distinguidos pelo tempo de chegada do fóton no segundo acoplador. Assim, escrevemos o estado do fóton logo após o segundo acoplador como:

$$|\Psi\rangle = c_1|s\rangle - c_2 e^{j\phi}|\ell\rangle \tag{2.22}$$

onde ϕ é a fase inserida pelo modulador de fase PM e os estados $|s\rangle, |\ell\rangle$ são assim chamados para representar os braços “curto” (*short*) e “longo” (*long*) do interferômetro, isto é, o “time bin” no qual o fóton se encontra. Os coeficientes reais de acoplamento c_i são definidos unicamente pelo dispositivo. Em geral,

faz-se a associação natural $|0\rangle \equiv |s\rangle$ e $|1\rangle \equiv |\ell\rangle$. É desejável que o espaçamento entre os modos "short" e "long" seja suficientemente grande para que não haja interferência entre ambos na saída do interferômetro; isso pode ser feito usando-se uma diferença de comprimento entre os braços que seja maior que o comprimento de coerência dos fótons. Dado que podemos escrever $c_1 = \cos(\theta/2)$ e $c_2 = \sin(\theta/2)$, (2.22) equivale a (2.6), e portanto o espaço de Hilbert gerado é isomórfico ao espaço de estados de polarização.

O qubit de (2.22), apesar de equivalente a um qubit codificado em polarização, é muito mais curioso de um ponto de vista conceitual. Na saída do interferômetro, ele pode ser visualizado como uma superposição coerente de dois pulsos espaçados pela diferença de comprimento entre os braços do interferômetro, como se o fóton tivesse, de fato, escolhido os dois caminhos ao mesmo tempo. É muito comum, na prática, escolher $c_1 = c_2 = 1/\sqrt{2}$ (ou seja, usar um acoplador 50/50 na entrada) [15] e um switch no lugar do segundo acoplador.

Para realizarmos uma medida no qubit descrito por (2.22) na base computacional, basta simplesmente medir-se o tempo de chegada do fóton. No entanto, caso deseje-se medir em outra base, basta utilizar um aparato idêntico ao utilizado para gerá-lo, com a diferença que, nas duas saídas do segundo acoplador, inserimos dois detectores. Mais uma vez, a escolha dos coeficientes de acoplamento e da fase do modulador determinam a base na qual a medida é feita.

2.2. Geração de fótons únicos

Vimos que, em sistemas de comunicações quânticas, a informação é codificada em fótons únicos. Uma fonte ideal seria capaz de produzir, a cada ciclo de repetição, pulsos de luz contendo apenas um fóton, isto é, no estado de Fock $|1\rangle$. A pergunta que temos que nos fazer agora é: como, de um ponto de vista tecnológico, podemos fazer para obter uma fonte desse tipo?

[15] Nesse caso, na representação na esfera de Bloch, apenas os estados pertencentes a um grande círculo que passa pelos pontos $|s\rangle$ e $|\ell\rangle$ podem ser gerados por este método.

Na realidade, esse é um assunto completamente não-trivial. Fontes de fótons isolados, também conhecidas como *photon guns* [16], ainda estão longe de se tornar realidade. O que se faz na prática é trabalhar com lasers atenuados, de forma que o número médio de fótons por pulso seja tão pequeno que a probabilidade de haver mais de um fóton no mesmo pulso seja tão pequena quanto se queira.

2.2.1. Lasers atenuados

A forma mais usual de se gerar fótons únicos consiste na utilização de lasers altamente atenuados. No entanto, a luz de um laser não possui fótons igualmente espaçados no tempo, de forma que seja possível que todos os pulsos contenham um único fóton. Na realidade, os fótons produzidos por um laser se encontram em um *estado coerente*, que é uma superposição dos estados de Fock dada por:

$$|\mu\rangle = e^{-\mu/2} \sum_n \frac{\sqrt{\mu}^n}{\sqrt{n!}} |n\rangle \tag{2.23}$$

Na expressão acima, μ é o *número médio de fótons por pulso*. Não é muito difícil perceber que estados coerentes são a melhor descrição quântica para a luz de um laser. Para entendermos esse processo, vamos tomar um feixe de luz clássica como ponto de partida.

Assim, considere um laser emitindo um feixe de luz de potência constante P. Sabemos que o fluxo médio de fótons, em fótons por segundo, emitidos por um laser quasi-monocromático é dado por $\Phi = P/\hbar\omega$, onde ω é a freqüência óptica. Desta forma, em um dado intervalo de tempo τ (correspondente à duração de um pulso), podemos afirmar que aproximadamente $\alpha\Phi\tau$ fótons foram capazes de atravessar o atenuador, cujo coeficiente de transmissão é α.

Vamos agora dividir o intervalo de tempo τ em N sub-intervalos de comprimento τ/N, de forma que não haja mais de um fóton em um dado sub-intervalo. Dessa forma, cada intervalo tem uma probabilidade $p = \alpha\Phi\tau/N$ de

possuir um fóton e probabilidade $1-p$ de estar vazio. A probabilidade de se encontrar n fótons distribuídos pelos N intervalos é a mesma probabilidade de uma moeda viciada dar "cara" em N lançamentos – ou seja, o problema segue uma distribuição *binomial*, que é dada por:

$$\begin{aligned} p(n) &= \binom{N}{n} p^n (1-p)^{N-n} = \frac{N!}{n!(N-n)!}\left(\frac{\alpha\Phi\tau}{N}\right)^n \left(1-\frac{\alpha\Phi\tau}{N}\right)^{N-n} \\ &= \frac{(\alpha\Phi\tau)^n}{n!}\left\{\frac{N!}{(N-n)!N^n}\right\}\left(1-\frac{\alpha\Phi\tau}{N}\right)^{N-n} \end{aligned} \tag{2.24}$$

Tomando o limite quando $N \to \infty$, o termo entre chaves tende a 1 e o último termo tende a $\exp(-\alpha\Phi\tau)$, de forma que obtemos

$$p(n) = \frac{\mu^n}{n!} e^{-\mu} \tag{2.25}$$

onde $\mu = \alpha\Phi\tau$. Observe que (2.25) é exatamente a distribuição de probabilidade obtida pela expressão (2.23): basta calcular $|\langle n|\mu\rangle|^2$ e verificar que o resultado é o mesmo.

Note que (2.25) é uma distribuição de probabilidade poissoniana e, por esta razão, não é possível obter uma quantidade indefinida de pulsos consecutivos que contenham exatamente um fóton. Por motivos que se tornarão claros na seção 2.5 sob o contexto de criptografia quântica, não é desejável gerar-se pulsos contendo mais de um fóton. Para que isso seja evitado, observe que a probabilidade de um pulso não-vazio possuir mais de um fóton é dada por:

$$\begin{aligned} \mathrm{P}(n>1 \,|\, n>0) &= \frac{1-p(0)-p(1)}{1-p(0)} \\ &= \frac{1-e^{-\mu}(1+\mu)}{1-e^{-\mu}} \approx \frac{\mu}{2} \end{aligned} \tag{2.26}$$

Logo, a probabilidade de haver pulsos multi-fóton pode ser feita tão pequena quanto se queira. É claro que isso tem um preço: quanto menor o valor

médio de fótons por pulso, maior a probabilidade de emissão de pulsos vazios, dada por $p(0) = e^{-\mu} \approx 1 - \mu$. Portanto, o problema dos lasers atenuados está na redução da taxa de transmissão que ocorre quando μ é feito muito pequeno. Esse problema, a princípio, poderia ser eliminado se a freqüência de transmissão fosse aumentada – de MHz para GHz, por exemplo – mas, como veremos na seção 2.4, o ruído na detecção não se altera e, assim, a relação sinal-ruído diminui.

2.2.2. Fonte de fótons anunciados

Uma forma engenhosa de evitar o efeito negativo causado pelo aumento da probabilidade de emissão de pulsos vazios foi introduzida por Hong e Mandel [17]. A figura 4, abaixo, ilustra a idéia, que envolve o uso de um cristal não-linear.

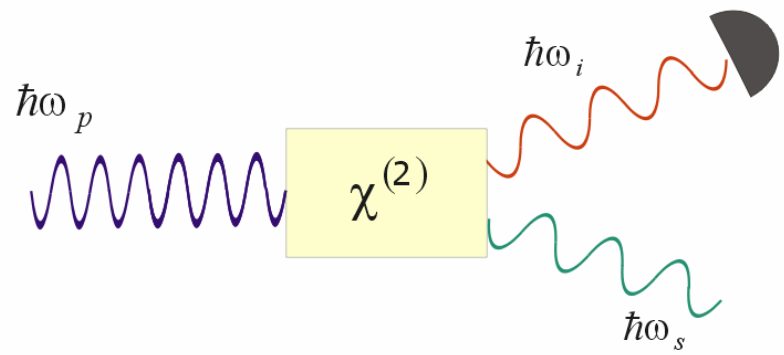

Figura 4. Esquema simplificado para fonte de fótons "anunciados". O fóton *idler* é utilizado como um testemunho da existência do fóton *signal*.

No esquema proposto, um laser de bombeio incide sobre um cristal não-linear de segunda ordem de forma a gerar pares de fótons pelo processo de *spontaneous parametric down-conversion*[16]. Devido à conservação da energia, a existência de um fóton na freqüência ω_s ("*signal*") implica na existência de outro na freqüência ω_i ("*idler*") criado exatamente ao mesmo tempo, satisfazendo $\omega_p = \omega_s + \omega_i$ onde ω_p ("*pump*") é a freqüência do laser de bombeio. Portanto, o fóton em ω_i pode servir de "testemunho" para a existência do fóton que está

[16] Esse processo é o exato oposto da geração de soma de freqüências, discutida em detalhes no capítulo 3, e é muito utilizado para a geração de pares de fótons emaranhados.

sendo enviado. O transmissor pode, desta forma, avisar o receptor, via comunicação clássica, em quais momentos um pulso foi realmente enviado.

Para que o esquema funcione da melhor forma possível, é necessário que a eficiência do processo não-linear seja a mais alta possível (em geral, ela não passa de $\approx 10^{-6}$). Além disso, o detector utilizado para "anunciar" a geração de um fóton deve ter eficiência quântica suficientemente alta, caso contrário apenas uma pequena fração dos pulsos contendo fótons serão anunciados.

Observe que esse esquema *não* soluciona o problema de pulsos multi-fóton, pois, visto que o laser utilizado para bombeio do cristal não-linear precisa ser muito intenso, existe uma probabilidade de que dois pares de fótons sejam criados ao mesmo tempo.

2.3. O canal quântico

Assim como em qualquer sistema de telecomunicações, a fonte e o receptor se comunicam através de um canal. No caso de comunicações quânticas, ele é chamado de *canal quântico*.

Apesar do nome, o canal em si não possui nada de "quântico" no sentido usual da palavra, pois trata-se do mesmo canal utilizado também para comunicações clássicas. Entretanto, visto que muitas aplicações de comunicações quânticas também envolvem uma comunicação clássica entre transmissor e receptor – como os exemplos da criptografia quântica e do teletransporte quântico – criou-se uma distinção entre os dois canais. Ou seja, o nome "canal quântico" é utilizado simplesmente para enfatizar o fato de que os portadores de informação que o atravessam são entidades quânticas, diferentemente do que ocorre no *canal clássico*. Ora, então qual é a diferença entre os dois – se é que há alguma – além da natureza do portador[17]?

Há muitas diferenças! Mesmo que ambos os canais, clássico e quântico, compartilhem espaços *físicos* idênticos, a caracterização do canal se dá de forma bem distinta. Um bom exemplo para ajudar a compreender a distinção conceitual entre os canais é o efeito de *atenuação*. Sabe-se que, em qualquer canal real, o

[17] Na discussão que se segue, supõe-se um sistema digital para as comunicações clássicas.

sinal está sujeito a sofrer atenuação por um ou vários fatores; todavia, enquanto se fala sobre "pulsos atenuados" em comunicações clássicas – que podem resultar em uma decisão equivocada por parte do receptor em qual bit foi transmitido – não podemos falar em "fótons atenuados" no contexto de comunicações quânticas. Ou ele chega ao receptor, ou não chega – não há nenhuma opção intermediária nesse sentido, por definição. Assim como muitos conceitos em física quântica, a atenuação surge como um efeito *estatístico* que só pode ser verificado quando o mesmo experimento é realizado muitas vezes, isto é, ela retrata a *probabilidade* de um fóton conseguir atravessar o canal sem ser destruído.

Como, portanto, modelar um canal quântico? Muitos trabalhos já foram realizados nessa área, e para responder a essa pergunta, utiliza-se a chamada *teoria dos sistemas quânticos abertos*, que trata a questão dos efeitos de um ambiente incontrolável na evolução de um sistema quântico. A figura 5 ilustra a diferença entre sistemas quânticos abertos e fechados.

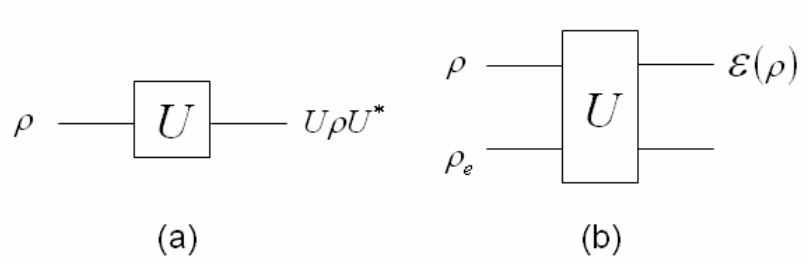

Figura 5. (a) Evolução de um sistema quântico isolado; (b) Evolução de um sistema quântico aberto, no qual leva-se em conta a interação com o ambiente.

Uma forma natural de se modelar um sistema quântico aberto é considerar que o sistema original, chamado de *principal*, interage com um outro sistema quântico chamado de *ambiente*. Nessa modelagem, os sistemas principal e ambiente, juntos, foram um sistema fechado[18]. Assumindo que, inicialmente, o estado global do sistema é um estado produto, o estado final do sistema principal, após propagação pelo canal quântico, é dado por:

[18] O termo *ambiente*, nesse contexto, representa qualquer outro grau de liberdade que não seja aquele no qual a informação é codificada, e portanto não necessariamente se refere a algo externo ao fóton.

$$\mathcal{E}(\rho) = \mathrm{tr}_e \left[U(\rho \otimes \rho_e) U^* \right] \tag{2.27}$$

Nessa notação, ρ_e representa o estado do ambiente, sobre o qual é tomado o traço parcial. Em qualquer situação, é desejável que o canal mantenha isolado o grau de liberdade no qual a informação foi codificada; em outras palavras, a transformação introduzida pela propagação de um qubit pelo canal quântico deve, idealmente, provocar uma evolução do tipo

$$\rho \otimes \rho_e \xrightarrow{U} \tilde{U}\rho \otimes \tilde{U}_e \rho_e \tag{2.28}$$

isto é, $U = \tilde{U} \otimes \tilde{U}_e$. Nesse caso, $\mathcal{E}(\rho) = \tilde{U}\rho\tilde{U}^*$ e o sistema principal permanece isolado do ambiente. Note que, mesmo com uma situação desse tipo, ainda pode haver problemas! Isso ocorre pois, em geral, a transformação U pode variar com o tempo de forma aleatória (pense em uma trajetória aleatória na esfera de Bloch). Portanto, no caso ideal, essa transformação é invariante no tempo, de forma que seu efeito possa ser compensado no receptor pela aplicação da transformação inversa $\tilde{U}^*$.

Se (2.28) não for satisfeita, significa que houve uma interação entre sistema e ambiente. Um ponto de vista complementar argumenta que uma interação que não satisfaça (2.16) constitui um *vazamento de informação* do sistema para o ambiente, no sentido que uma medida realizada no ambiente pode revelar propriedades do qubit original [18]. Do ponto de vista da criptografia quântica, por exemplo, isso é certamente indesejável.

Escolhendo-se uma transformação unitária U adequada, pode-se modelar todos os efeitos que um canal pode causar em um qubit, tais como atenuação e introdução de ruído[19], além de um tipo de perturbação sem análogo nos sistemas de comunicação clássica, chamado de *descoerência*, no qual há perda de

[19] Na realidade, atenuação e adição de ruído podem ser modelados simultaneamente considerando que o ambiente atua como um separador de feixes. No caso da atenuação, um fóton do sistema é perdido para o ambiente, e no caso do ruído, um fóton do ambiente é "perdido" para o sistema.

informação (isto é, perda da fase relativa entre elementos de uma base para o qubit) sem que haja perda de energia.

Na prática, o canal quântico poderia consistir em qualquer meio em que a luz possa se propagar, mas as escolhas mais apropriadas são as mesmas dos sistemas de comunicações ópticas clássicos: as *fibras ópticas* e a *atmosfera*. Intuitivamente, poderíamos dizer que fibras ópticas são mais próximas do ideal de "sistema fechado" pelo fato do sinal se encontrar espacialmente confinado, enquanto a atmosfera é aparentemente mais "aberta" a influências externas. Entretanto, essa intuição é equivocada! Alguns tipos de fibra, chamadas de *fibras multi-modo*, admitem a propagação de vários modos espaciais do sinal luminoso ao mesmo tempo, que se acoplam facilmente entre si e, assim, atuam no qubit como um ambiente não-isolado. A intuição está correta, no entanto, para *fibras monomodo*, que guiam apenas um modo de luz e por isso são muito apropriadas como canal quântico. Embora a descrição de fibras ópticas como canal quântico não pertença ao escopo dessa tese, o capítulo 5 contém muitas informações sobre o canal atmosférico.

2.4. Detecção de fótons únicos

A existência de (pseudo-) fontes de fótons únicos e de canais através dos quais a informação quântica pode ser transmitida de nada adiantariam se não houvesse meios de *detectar* pulsos de luz contendo apenas um fóton. De fato, a tecnologia existe há algum tempo – embora não para qualquer comprimento de onda – e pode ser implementada de diversas formas, tais como tubos fotomultiplicadores, junções supercondutoras ou fotodiodos avalanche (APDs). Esses dispositivos são denominados *módulos de contagem de fótons únicos* (SPCM, *single photon counting module*), ou simplesmente *contadores de fótons*. Vale lembrar que esses dispositivos têm sido amplamente empregados em outras áreas de pesquisa, tais como metrologia, astronomia e espectroscopia.

Esta seção está organizada da seguinte forma: inicialmente, os parâmetros relevantes e figuras de mérito para comparação de contadores de fótons são introduzidas, tomando os APDs como exemplo. Em seguida, discutimos o problema de detecção de um ponto de vista prático, com relação ao comprimento

de onda dos fótons, e fornecendo algumas soluções alternativas (ou mandatórias) ao uso de APDs.

2.4.1. Caracterização de um contador de fótons

Quais seriam as características de um contador de fótons ideal? Antes de responder a essa pergunta, é preciso identificar as propriedades dos detectores que são relevantes no problema de contagem de fótons. São eles: *eficiência quântica*, *probabilidade de ruído*, *resolução temporal* e *tempo morto*.

Na discussão a seguir, consideramos que o contador de fótons utiliza a tecnologia dos APDs, que além de serem os detectores mais largamente utilizados em aplicações práticas, foram também os utilizados no trabalho experimental dessa tese[20]. Consideramos APDs de *silício*, que são sensíveis à faixa espectral 400-1000 nm, e APDs de *InGaAs/InP* (arseneto de índio-gálio), utilizados para contagem de fótons no comprimento de onda de telecomunicações 1.55μm.

2.4.1.1. Eficiência quântica

A *eficiência quântica* de um contador de fótons é definida como a probabilidade que um fóton incidente gere um pulso[21] elétrico. No caso dos APDs, essa eficiência é um produto de três eficiências primárias: em primeiro lugar, o fóton deve atingir a região fotosensível do detector; em seguida, deve ser absorvido, de forma a gerar um par elétron-buraco e, por último, esse par elétron-buraco deve ser capaz de iniciar uma avalanche. Isso pode ser resumido pela expressão

$$\eta_{\text{det}} = (1-R)\left(1-e^{-\alpha d}\right)\eta_{av} \tag{2.29}$$

[20] Para uma descrição do funcionamento de APDs, ver [19].

[21] O termo "pulso" é utilizado tanto para descrever o sinal elétrico gerado na detecção de um fóton como para se referir à duração de um pulso do laser utilizado na geração do sinal.

onde R é o coeficiente de reflexão na superfície, α é o coeficiente de absorção do material (ambos, em geral, dependentes do comprimento de onda da luz incidente) e η_{av} a probabilidade de geração de avalanche (dependente da eletrônica utilizada).

Para um contador de fótons ideal, a eficiência quântica deve ser a mais próxima possível de 100%, ao longo da maior faixa espectral possível, mas raramente se consegue na prática um valor acima de 75% para APDs de silício a 780 nm ou 40% para APDs de InGaAs a 1550 nm. O efeito de uma baixa eficiência quântica é exatamente o mesmo efeito causado por perdas (atenuação) no canal, e pode tornar um sistema de comunicações quânticas inviável, mesmo que a relação sinal-ruído seja alta.

2.4.1.2. Ruído

A *probabilidade de ruído* é definida como a probabilidade de um sinal elétrico ser gerado espontaneamente, independentemente da presença de fótons. Esse ruído surge do fato que pares elétron-buraco podem ser gerados por mecanismos outros que a absorção de um fóton, tais como processos de tunelamento entre as bandas de condução e valência ou, na maior parte dos casos, processos oriundos de agitação térmica. Visto que esse ruído ocorre mesmo sem luz incidindo no detector, ele é chamado de *ruído de escuro*, e os pulsos elétricos gerados de *contagens de escuro* (*dark counts*).

O ruído de escuro é um processo sem memória e pode ser modelado por uma variável aleatória poissoniana, exatamente da mesma forma que a geração de fótons por um laser atenuado. O termo análogo ao fluxo de fótons Φ é a *taxa de escuro* n_{dark}, que expressa o valor médio de contagens por unidade de tempo. Porém, não faz sentido falar em $p(n)$, ou seja, na probabilidade de haver n contagens de escuro por pulso, pois a cada "pulso" (ou seja, cada janela temporal de detecção), só pode haver, no máximo, uma contagem! Portanto, todos os casos onde $n > 0$ devem ser considerados de forma conjunta, de forma que

$$P(ruído) = 1 - p(0) = 1 - e^{-\mu} \tag{2.30}$$

onde $\mu = n_{dark} T$ é o número médio de contagens por pulso e $P(ruído)$ é a probabilidade de ruído *por pulso*. Na prática, a duração T dos pulsos é muito pequena (ordem de nanosegundos), de forma que μ é muito pequeno e (2.30) pode ser aproximada por $P(ruído) \approx \mu$.

Uma outra fonte de ruído em APDs são os chamados *afterpulses*, que estão associados ao efeito de cargas presas em "armadilhas" (níveis energéticos no interior do gap) devido a avalanches anteriores. Esse efeito pode ser reduzido aumentando-se o tempo morto – o que não é desejável, como veremos na próxima sub-seção – ou aumentando-se a temperatura, o que aumentaria a taxa de escuro e, portanto, também não é desejável. É devido aos *afterpulses* que não podemos resfriar um APD a temperaturas baixíssimas de forma a não obter contagens de escuro.

É evidente, portanto, que o contador de fótons ideal deve ser isento de ruído. Embora já haja módulos contadores de fótons de silício com ruído baixíssimo, exibindo $n_{dark} \approx 1\,\text{Hz}$, nada parecido foi obtido com APDs de InGaAs.

2.4.1.3. Resolução temporal e tempo morto

A *resolução temporal* de um detector é a incerteza no intervalo de tempo entre a detecção de um fóton e a geração de um pulso elétrico. Qualquer detector semicondutor possui um certo tempo de resposta devido a efeitos de difusão ou à capacitância do fotodiodo, mas no caso específico de APDs, há de se considerar o *tempo de construção da avalanche*, que é um processo aleatório (decorrente da aleatoriedade do processo de multiplicação por avalanche). É crucial que a resolução temporal seja suficientemente inferior à duração do pulso, de forma que o efeito de *jitter* seja desprezível. No caso de APDs de silício, resoluções temporais inferiores a 100 ps podem ser obtidas [20].

O *tempo morto* (ou tempo de recuperação) é o intervalo de tempo no qual o detector se "recupera" após o último pulso elétrico gerado. Um contador de fótons ideal não possui tempo morto, e a duração do intervalo entre duas janelas de detecção dependeria apenas de quão rapidamente o laser, no transmissor, seria pulsado; em sistemas reais, essa propriedade é conseqüência da eletrônica utilizada, e não apenas do fotodiodo em si. Em geral, APDs são operados no

chamado *modo Geiger*, no qual a tensão elétrica aplicada é superior à tensão de ruptura. Desta forma, um simples fóton é capaz de desencadear uma avalanche, composta de milhares de pares elétron-buraco. Uma vez desencadeada, o APD necessita de um meio de suprimi-la para que ele possa se preparar para receber um novo fóton. Esse processo de supressão da corrente macroscópica gerada é chamado de *quenching*.

A forma mais simples de se realizar essa operação consiste em interromper a avalanche de forma passiva (*passive quenching*). Neste caso, um resistor é conectado em série com o diodo de forma a causar uma diminuição em sua voltagem assim que a avalanche acontecer, de forma a reduzi-la abaixo da tensão de ruptura e reinicializar o APD. O tempo morto, nesse caso, será dado pela constante *RC* do circuito, de algumas centenas de nanosegundos. No entanto, um método sugerido por S. Cova *et al* [21] permitiu taxas de contagem mais elevadas. O procedimento utiliza uma eletrônica mais sofisticada para fazer uma detecção ativa da presença de avalanche (*active quenching*), sendo capaz de reduzir o tempo morto para alguns nanosegundos.

Se o tempo de chegada dos fótons for conhecido, como acontece em sistemas síncronos, também é possível operar o APD em um modo no qual a tensão é aumentada além do limiar de ruptura apenas para os instantes de tempo nos quais é esperada a chegada de um pulso. Esse método (*gated mode*) permite tempos mortos da mesma ordem de grandeza que os obtidos por *active quenching* com uma eletrônica mais simples, porém às custas de um sistema de sincronismo.

2.4.1.4. Critérios de comparação

Apesar das quatro propriedades apresentadas até agora serem suficientes para a caracterização de um contador de fótons, nem sempre são as mais utilizadas. Muitas vezes, é conveniente expressar várias características de forma compacta, a fim de facilitar a comparação entre diferentes detectores.

Em sistemas de comunicações clássicos, um conceito amplamente utilizado, e que pode ser trazido para as comunicações quânticas, é o de *relação sinal-ruído*. Já que trabalhamos constantemente no limiar permitido pelas leis da física, que é a detecção da energia de um único fóton, temos a falsa ilusão de que basta possuir

uma boa eficiência quântica para se obter um bom detector. No entanto, ser capaz de detectar um fóton não significa ser capaz de extrair informação a partir desse fóton, devido à presença de ruído na detecção. Essa relação próxima entre eficiência e ruído deu origem a algumas figuras de mérito – todas importadas dos sistemas clássicos.

A mais utilizada de todas se chama *potência equivalente de ruído* e abreviada como NEP (*noise equivalent power*). Para contadores de fótons, ela é expressa como:

$$NEP = \frac{hc}{\eta_{\text{det}}\lambda}\sqrt{2\langle n_{dark}\rangle} \tag{2.31}$$

Observe que a NEP depende do comprimento de onda do fóton (isto é, de sua energia), da eficiência quântica e do valor médio da taxa de ruído. A grande deficiência dessa figura de mérito é que ela não possui nenhuma interpretação física evidente, já que é medida em Watts por raiz quadrada de Hertz. Isso pode fazer sentido para detectores clássicos, nos quais o ruído tem uma dependência quadrática com a banda passante do detector, mas essa dependência não existe em contadores de fótons e (2.31) torna-se desprovida de significado.

Para solucionar esse problema, utilizamos ao longo dessa tese uma outra figura de mérito extremamente semelhante, a *sensibilidade*, que é dada por:

$$SNR_0 = \frac{hc}{\eta_{\text{det}}\lambda}\langle n_{dark}\rangle \tag{2.32}$$

A diferença é sutil, mas (2.32) possui uma interpretação física muito simples: a sensibilidade é a potência incidente no detector que corresponde a uma relação sinal-ruído igual a 1. Portanto, ela pode ser encarada como uma espécie de "limiar" de detecção.

Com relação à resolução temporal e ao tempo morto, costuma-se introduzir a figura de mérito *taxa máxima de contagem*, que corresponde à máxima taxa de

pulsos por segundo que poderia ser detectada sem sobreposição de pulsos vizinhos[22]. Ela é dada por

$$R_{\max} \cong \frac{1}{\tau_r + \tau_m} \tag{2.33}$$

onde os termos no denominador são a resolução temporal e o tempo morto. Lembre-se que esses termos são médias estatísticas.

Atente para a diferença entre essa taxa e a *taxa de dados*, que também depende da eficiência quântica e do tipo de protocolo de comunicação utilizado.

2.4.2. A dependência do comprimento de onda

Comentamos na seção anterior que a eficiência quântica η_{det} varia em função do comprimento de onda da luz que desejamos detectar. Na realidade, cada material (seja semicondutor, supercondutor, etc) possui um tipo de dependência distinto, sendo alguns mais sensíveis para certas regiões do espectro do que outros; portanto, a escolha do comprimento de onda de um sistema de comunicações quânticas vai depender essencialmente dos contadores de fótons disponíveis no mercado.

Mas quais são os comprimentos de onda mais adequados para um sistema de comunicações quânticas? De forma não muito surpreendente, são os mesmos comprimentos de onda dos sistemas de comunicações clássicos, afinal, já existe toda uma tecnologia disponível que pode ser imediatamente aproveitada[23]. Valores muito comuns são 780 nm, 850 nm, 1.3 μm e 1.55 μm.

Para qualquer comprimento de onda inferior a ~1 μm, o melhor método de contagem de fótons que existe é, sem sombra de dúvida, baseado em APDs de silício. Há mais de duas décadas um esforço considerável tem sido aplicado no desenvolvimento de módulos de contagem utilizando esses APDs, de forma que

[22] Alguns autores utilizam a taxa máxima simplesmente como o inverso do tempo morto, desconsiderando o problema do *overlap* entre pulsos.

[23] O capítulo 5 discute comprimentos de onda alternativos no caso de transmissão atmosférica dos qubits.

eles são considerados ideais para a tarefa, já tendo ultrapassado a antiga tecnologia de tubos fotomultiplicadores. Módulos comercialmente disponíveis são capazes de exibir eficiências quânticas superiores a 70% (para a faixa de comprimentos de onda 700-800 nm), resoluções temporais inferiores a 100 ps, taxas de contagem máxima de 10 MHz e, a temperaturas em torno de -50ºC (facilmente obtidas com o uso de um Peltier), taxas de escuro da ordem de 1 Hz! É claro que nem todas essas características formidáveis são encontradas ao mesmo tempo no mesmo dispositivo, mas de qualquer forma, APDs de silício são uma excelente solução para comunicações quânticas. Infelizmente, não há fibras ópticas capazes de provocar atenuações suficientemente baixas em fótons de comprimento de onda inferior a 1 μm, de forma que o uso de APDs de silício é restrito a sistemas de espaço livre.

Para comprimentos de onda acima de ~1 μm, no entanto, não há solução definitiva. Na segunda janela de telecomunicações (em torno de 1.3 μm), o problema tem sido tradicionalmente resolvido pelo uso de APDs de germânio (Ge) ou de arseneto de índio-gálio (InGaAs/InP). O problema principal dos APDs Ge é a necessidade de baixíssimas temperaturas, em torno de 77K; no entanto, taxas de escuro e eficiências aceitáveis (25 kHz e 10%, respectivamente) podem ser obtidas. Já na terceira janela de telecomunicações, apenas APDs de InGaAs podem ser utilizados. Apesar de não necessitar de temperaturas tão baixas (mas, mesmo assim, que não passam de 173K), o APD InGaAs possui uma maior probabilidade de *afterpulses*. No comprimento de onda de 1.55 μm, dispositivos com eficiência quântica de 25% e taxas de ruído de ~ 100 kHz já estão comercialmente disponíveis. Observe a enorme diferença entre APDs InGaAs e Si.

Devido à relativamente baixa performance de detectores InGaAs disponíveis no mercado, métodos de detecção alternativos foram sugeridos. Um deles consiste no uso da tecnologia TES (*transition edge sensor*), que consiste em microcalorímetros supercondutores originalmente desenvolvidos para aplicações em astronomia [22]. Os resultados são realmente incríveis: eficiência quântica de 20% e taxa de escuro de 0,01 Hz! No entanto, a temperatura crítica do material supercondutor utilizado é de 175 *mK*, o que é uma enorme desvantagem. Outra abordagem para o problema de contagem de fótons a 1.55 μm é o método de

conversão de freqüências, o tema central dessa tese. Esse assunto será detalhadamente abordado nos capítulos 4 e 5.

2.5. Criptografia quântica

A palavra *criptografia* é originada da combinação dos termos gregos κρυπτοσ (*kryptós*) "escondido" e γραφο (*gráfo*) "grafar" ou "escrever". Trata-se da arte de tornar uma mensagem ininteligível a qualquer um que não seja o destino correto, cuja raiz data dos primórdios de nossa civilização. Vestígios históricos mostram que Julio César desenvolveu uma cifra de substituição para se comunicar secretamente com terras distantes do império romano, e também é de nosso conhecimento que a criptografia militar já era utilizada pelos espartanos em suas famosas guerras contra os gregos. Da mesma forma, outras civilizações da África e Ásia também desenvolveram técnicas de transmitir informações confidenciais.

Nos dias de hoje, a criptografia deixou de ser um luxo para se tornar uma necessidade. Virtualmente todas as transações financeiras e comerciais da atualidade dependem, de alguma forma, de poderosos algoritmos de criptografia, sem os quais o mundo definitivamente não seria o mesmo. Todos nós, quando compramos um livro pela Internet usando nosso número de cartão de crédito, estamos, ao menos inconscientemente, confiando na segurança que nos é dada por esses algoritmos. A última coisa que desejamos, afinal, é que o número do nosso cartão caia em mãos erradas.

Mas será que estamos verdadeiramente seguros? Até que ponto podemos confiar nos esquemas de criptografia tão amplamente utilizados? Curiosamente, a resposta é que *não se sabe*. De fato, os sistemas de criptografia utilizados na maioria das aplicações comerciais são os chamados *sistemas de chave pública* ou *sistemas assimétricos*, nos quais a segurança é baseada em complexidade computacional. Entenda por *complexa* uma operação "difícil" para a qual seriam necessários métodos de força bruta que, com o poder computacional existente hoje em dia, poderia levar anos ou mesmo décadas para ser realizada. Um bom exemplo, e que é largamente utilizado em protocolos de criptografia, é o problema de fatoração de números muito grandes, como por exemplo fatorar um número de

617 dígitos em seus dois fatores primos[24]. Embora até os dias de hoje não exista nenhum método para realizar tais tarefas, elas não são *a princípio* impossíveis. Mesmo que a probabilidade de alguém surgir, do dia pra noite, com um algoritmo de fatoração rápida seja praticamente nula, não podemos nos esquecer que o poder computacional dos processadores está aumentando progressivamente, e com o advento do *computador quântico*, o ganho de poder de processamento será *exponencial*. Não é muito difícil perceber que, cedo ou tarde, a criptografia de chave pública estará com seus dias contados.

A solução reside na utilização de sistemas de chave *simétrica*. Sistemas simétricos são aqueles nos quais uma mesma chave é utilizada nos processos de codificação (*encriptação*) e decodificação (*decriptação*). Uma analogia a esse sistema é um cofre no qual uma mensagem é trancada e enviada ao destinatário; para recuperar a mensagem, o cofre deve ser aberto com a mesma chave com que foi trancado. Já foi demonstrado que existem sistemas simétricos *incondicional-mente seguros*, isto é, que não podem ser quebrados sem o conhecimento da chave.

Ora, por que não utilizar um sistema desse tipo no lugar dos sistemas assimétricos? E o que a física quântica tem a ver com essa história toda? Essas e outras perguntas serão respondidas nas sub-seções a seguir.

2.5.1.
O problema da distribuição de chaves

Antes de prosseguirmos, vamos apresentar os personagens que geralmente surgem nos textos tradicionais de criptografia: Alice e Bob, dois indivíduos que desejam se comunicar secretamente, e Eva, uma espiã. Nesse cenário, Alice e Bob desejam compartilhar uma chave secreta para trocar informações e Eva deseja obter a maior informação possível sobre essa chave. Na maioria dos casos, vamos supor que Alice deseja enviar uma mensagem a Bob.

Prosseguindo a discussão, foi afirmado que sistemas de chave simétrica são seguros. Para demonstrarmos esse fato, vamos utilizar como exemplo a famosa cifra proposta por Gilbert Vernam em 1926, chamada de *one-time pad*. Nesse

[24] Para os interessados, esse problema ainda está em aberto e a empresa RSA oferece 200 mil dólares a quem conseguir solucioná-lo!

esquema, Alice encripta sua mensagem, uma seqüência de bits representada pelo número (binário) *m*, utilizando uma chave aleatória *k*. O procedimento é muito simples: Alice soma cada bit da mensagem ao bit correspondente da chave para obter o criptograma. Essa operação pode ser representada como:

$$s = m \oplus k \tag{2.34}$$

onde o símbolo $\oplus$ representa adição módulo 2 (ou, se preferir, um *ou exclusivo* bit a bit). O criptograma *s* é então enviado a Bob, que decripta a mensagem exatamente da mesma forma (daí a denominação "simétrico"), obtendo a seqüência de bits $\hat{m}$ dada por:

$$\hat{m} = s \oplus k = m \oplus k \oplus k = m \tag{2.35}$$

Ou seja, a mensagem original é recuperada. Como os bits do criptograma são tão aleatórios como a chave, eles não contêm nenhuma informação. Alguns anos depois, Shannon demonstrou que, enquanto (1) a chave fosse verdadeiramente aleatória, (2) tivesse o mesmo tamanho (em bits) que a mensagem e (3) nunca fosse reutilizada, então o one-time pad é *perfeitamente seguro*. Se temos um sistema perfeitamente seguro, cabe a pergunta: o que há, então, de errado com a criptografia clássica?

O problema se chama *distribuição de chaves*. Uma vez que Alice e Bob possuem uma chave secreta em comum, a comunicação subseqüente envolve o envio de criptogramas ao longo de um canal, que poderia inclusive ser um canal totalmente vulnerável a espionagem passiva (por exemplo, radiodifusão). Esse estágio é certamente seguro. Entretanto, para estabelecer a chave, é necessário que os dois usuários utilizem um canal muito seguro e confiável; e por mais seguro que seja, é sempre possível, a princípio, que esse canal possa ser passivamente monitorado, mesmo que isso seja extremamente difícil. Uma forma de contornar esse problema seria estabelecer uma chave de uma vez por todas, como por exemplo Alice contratar um carro-forte para enviar a chave para Bob. Como a chave deve ser renovada a cada mensagem, no entanto, o tamanho total das mensagens estará limitado ao tamanho da chave, o que torna esse processo proibitivamente caro. Por essa razão, na maioria das aplicações não se utiliza

processos absolutamente seguros; utiliza-se, na prática, sistemas menos caros e menos seguros, como os sistemas de chave assimétrica.

É exatamente aqui que entra a física quântica. A idéia da *criptografia quântica* (CQ) é realizar a distribuição de chaves usando – pasmem – um sistema de comunicações quânticas! Imagine o seguinte cenário: Alice gera números aleatórios ("0"s e "1"s) e os codifica em qubits, enviando-os para Bob através de um canal quântico, no qual há a presença da espiã Eva. Espionagem, do ponto de vista físico, é baseado em um conjunto de *medidas* realizadas pelo espião nos portadores de informação, que no caso são os fótons. Como sabemos da física quântica, se um dado conjunto de medidas não for compatível com o estado do sistema, então este será inexoravelmente perturbado – perturbações essas que podem ser detectadas posteriormente por Alice e Bob[25].

Dessa forma, se Alice e Bob utilizarem um sistema de comunicações quânticas para transmitir uma chave, temos duas alternativas: se não houver perturbações, as leis da física quântica nos garantem que ninguém teve acesso a essa chave e, portanto, ela foi transmitida com segurança; se houver perturbações, Alice e Bob sabem que alguém obteve informação a respeito da chave, e assim eles simplesmente a descartam! Como não há informação na chave, eles não perderam nada. Por essa razão, a criptografia quântica é mais corretamente chamada de *distribuição quântica de chaves* (QKD, *quantum key distribution*).

A sub-seção a seguir mostra de que forma Alice deve codificar a informação nos qubits de modo que qualquer medida de Eva resulte em perturbações no canal.

2.5.2. Codificação em bases não-ortogonais

Vamos imaginar a seguinte situação: Alice deseja transmitir bits para Bob por meio de um canal quântico. Suponha que os bits "0" e "1" são codificados segundo o esquema

---

[25] Basta lembrarmos do clássico experimento da fenda dupla, no qual um feixe de fótons (ou elétrons) passa através de duas fendas e atinge um anteparo, produzindo uma figura de interferência. Se tentarmos adivinhar por qual das fendas cada fóton passou, para nosso espanto, essa medida faz com que a figura de interferência desapareça. Von Neumann chamava esse fenômeno de "colapso da função de onda".

$$\begin{aligned} 0 &\mapsto |\phi\rangle \\ 1 &\mapsto |\varphi\rangle \end{aligned} \tag{2.36}$$

onde os qubits $|\phi\rangle, |\varphi\rangle$ são estados arbitrários distintos, que por conseguinte formam uma base para o espaço de Hilbert $\mathcal{H}$. Imagine agora que Eva deseja obter informações sobre os bits que estão sendo transmitidos. O que ela poderia fazer para cumprir essa tarefa? A tarefa mais simples que pode ser imaginada consiste em *fazer cópias dos qubits enviados*. Idealmente, isso poderia ser modelado como a aplicação de uma transformação unitária *U* que não perturbe os qubits sendo enviados mas que, ao mesmo tempo, forneça informação a Eva. No contexto dos canais quânticos introduzido na seção 2.3, Eva faria parte do "ambiente" e aplicaria uma transformação do tipo (2.28), na qual o sistema principal (no caso, os qubits enviados) não sofre qualquer influência do ambiente. Dado que Alice transmite estados puros, podemos representar a transformação ideal a ser introduzida por Eva como

$$|\psi^A\rangle \otimes |\psi_0^E\rangle \xrightarrow{U} |\psi^A\rangle \otimes |\psi^A\rangle \tag{2.37}$$

onde $|\psi_0^E\rangle$ é o estado inicial do sistema usado por Eva ("ambiente"). A transformação *U* de (2.37) representa uma *máquina de clonagem*.

Temos más notícias para Eva, no entanto: máquinas de clonagem como a de (2.37) ou similares *não podem existir*. Não é difícil demonstrar essa afirmação; basta escrevermos (2.37) substituindo o estado genérico enviado por Alice pelos dois estados $|\phi\rangle, |\varphi\rangle$ para obtermos o par de equações:

$$\begin{aligned} U\left(|\phi\rangle|\psi_0^E\rangle\right) &= |\phi\rangle|\phi\rangle \\ U\left(|\varphi\rangle|\psi_0^E\rangle\right) &= |\varphi\rangle|\varphi\rangle \end{aligned} \tag{2.38}$$

Tomando o produto interno das duas equações, obtemos:

$$U^*U\langle\varphi|\phi\rangle\langle\psi_0^E|\psi_0^E\rangle = \langle\varphi|\phi\rangle\langle\varphi|\phi\rangle \tag{2.39}$$

Isto é,

$$\langle\varphi|\phi\rangle = \langle\varphi|\phi\rangle^2 \tag{2.40}$$

Observe que ambos os lados de (2.40) são números complexos, e que a igualdade $x = x^2$ só possui solução se $x = 0$ ou $x = 1$. Portanto, (2.40) implica em $|\phi\rangle = |\varphi\rangle$ ou $|\phi\rangle \perp |\varphi\rangle$, isto é, uma máquina de clonagem quântica só é capaz de fazer cópias dos qubits enviados por Alice se (1) todos forem iguais entre si (e no caso não há transmissão de informação) ou (2) os qubits forem codificados em estados ortogonais entre si. Esse resultado é conhecido como o *teorema da não-clonagem*.

Portanto, Alice pode evitar que Eva faça cópias perfeitas de seus qubits desde que sejam codificados em estados não-ortogonais. Uma questão surge nesse ponto: o teorema da não-clonagem implica na impossibilidade de se distinguir dois estados não-ortogonais[26]. Ou seja, Bob não terá meios de dizer, de forma determinística, se o qubit enviado por Alice foi um "0" ou um "1", e portanto aparentemente frustrando a nossa tentativa de se estabelecer uma comunicação entre ambas as partes!

A solução para esse problema é muito simples e pode ser resolvida de diversas formas, os chamados *protocolos* de criptografia quântica. A seção a seguir descreve o primeiro protocolo desenvolvido, o que é suficiente para perfeita compreensão de como funciona a criptografia quântica.

[26] Pense na contrapositiva: se fosse possível distinguir estados não-ortogonais, bastaria que Eva realizasse medidas nos qubits enviados por Alice. Após descobrir em qual estado se encontram, poderia preparar novos estados exatamente idênticos aos que mediu e reenviá-los a Bob – o que, na prática, seria equivalente ao processo de clonagem.

2.5.3.
O protocolo BB84

No ano de 1984, em uma conferência da IEEE em Bangalore, India, os pesquisadores Charles Bennett e Gilles Brassard propuseram o primeiro protocolo para criptografia quântica que, hoje em dia, é chamado de BB84 [23]. É curioso constatar que esse trabalho pioneiro passou quase desapercebido na época, e que apenas uma década depois a comunidade científica começou a demonstrar os primeiros sinais de interesse.

O protocolo BB84 consiste no uso de 4 estados quânticos que constituam duas bases ortogonais distintas $\left(|u_1\rangle,|u_2\rangle\right)$ e $\left(|v_1\rangle,|v_2\rangle\right)$ que sejam *maximamente conjugadas,* isto é, que satisfaçam $\left|\langle v_i|u_j\rangle\right| = 1/\sqrt{2}$, como por exemplo as bases formada por $\left(|0\rangle,|1\rangle\right)$ e por $\left(|+\rangle,|-\rangle\right)$. Para cada uma das bases, um dos qubits representa um bit "0" e o outro um bit "1". Tradicionalmente, essa associação é feita da forma:

$$\begin{aligned} 0 &\mapsto |0\rangle \text{ ou } |+\rangle \\ 1 &\mapsto |1\rangle \text{ ou } |-\rangle \end{aligned} \tag{2.41}$$

Portanto, no ato de envio, Alice precisa realizar duas escolhas: a escolha de *qual bit* deseja enviar e a escolha de *em qual das duas bases* vai codificá-lo. Ambas as escolhas devem ser feitas de forma aleatória, da mesma forma que os bits compondo a cifra de Vernam precisam ser aleatórios. Estatisticamente, portanto, Alice escolherá a base $\left(|0\rangle,|1\rangle\right)$ em 50% do tempo e a base $\left(|+\rangle,|-\rangle\right)$ nos outros 50% do tempo, de forma que os quatro estados $\left(|0\rangle,|1\rangle,|+\rangle,|-\rangle\right)$ são enviados com a mesma probabilidade de 25% cada um. Quando Bob receber o qubit[27], ele independentemente faz outra escolha aleatória: em qual das duas bases vai medi-lo.

Por que o uso de bases maximamente conjugadas? Observe que se um qubit enviado por Alice for medido na base "errada" – por exemplo, se o qubit $|+\rangle$ (que

[27] Vamos inicialmente assumir que o canal quântico não introduz ruído.

codifica um “0”) for medido na base $\left(|0\rangle,|1\rangle\right)$ – as probabilidades de se obter “0” e “1” são dadas por:

$$\begin{aligned} p(0) &= \left|\langle 0|+\rangle\right|^2 = 1/2 \\ p(1) &= \left|\langle 1|+\rangle\right|^2 = 1/2 \end{aligned} \tag{2.42}$$

Ou seja, sempre que Bob escolher a mesma base com a qual Alice codificou o qubit, ele obterá o mesmo bit que foi codificado por Alice; no entanto, se a base “errada” for escolhida, existe uma probabilidade de 50% da medida resultar em um *erro*. Portanto, 25% dos bits recebidos por Bob conterão erros, uma taxa alta demais para qualquer tipo de comunicação.

Existe uma forma, no entanto, de identificar quais dos bits obtidos por Bob contêm erros. Bastaria que Alice e Bob também pudessem se comunicar de forma *clássica*, da forma representada na figura X. Dessa forma, para cada qubit recebido por Bob, ele poderia anunciar publicamente qual base utilizou para medir, e Alice responderia se ela usou ou não a mesma base para codificá-lo. Assim, em todas as ocasiões nas quais ambos escolheram a mesma base, eles compartilharão bits idênticos, que poderão compor a chave secreta; nas demais situações, os bits são descartados.

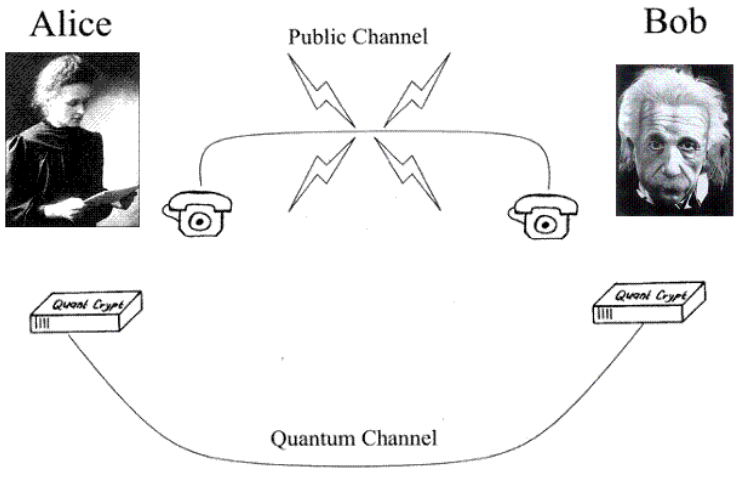

Figura 6. Esquema de um sistema de criptografia quântica, no qual Alice e Bob estão interligados por um canal quântico e um canal clássico.

É importante ressaltar que o canal clássico usado para reconciliação de bases seja *público*, ou seja, que possa ser monitorado por qualquer um mas impeça qualquer tipo de modificação na informação transmitida por ele. Radiodifusão é um bom exemplo desse tipo de canal. Observe ainda que, em momento algum, os valores dos bits são anunciados pelo canal clássico, mas apenas a *base* na qual a medida foi feita. De qualquer forma, temos a impressão de que essa informação poderia ser útil para Eva, e ela de fato pode ser. Visto que Eva não pode realizar cópias perfeitas dos qubits de acordo com o teorema da não-clonagem, ela pode partir para uma estratégia diferente. Um exemplo simples de estratégia de espionagem é o chamado método *interceptação-reenvio*.

Suponha que Eva se posiciona em algum ponto do canal quântico e, secretamente, realiza medidas nos qubits exatamente da mesma forma que Bob: escolhe aleatoriamente uma base de medida e obtém valores de bits. Em seguida, Eva prepara um qubit no mesmo estado que ela mediu e o envia para Bob, se fazendo passar por Alice. Como a reconciliação de bases é feita por um canal público, ela poderia obter informações sobre a chave em 50% dos bits, correspondendo às situações na qual Eva e Bob escolheram a mesma base de medida.

Por que a estratégia interceptação-reenvio não funciona? Observe que o estado reenviado por Eva pode não ser igual ao estado que ela recebeu de Alice, com uma probabilidade de 50%, pois ela pode escolher a base errada. Nesses casos, sempre que Bob escolher a mesma base que Alice, um erro será obtido – ou seja, em 25% dos bits da chave após reconciliação de bases. Essa altíssima taxa de erro pode ser facilmente detectada tomando amostras da seqüência e comparando resultados por meio do canal clássico. É exatamente nesse sentido que a presença do espião introduz "distúrbios" no sistema, que podem ser detectados.

É claro que Eva poderia realizar essa estratégia em apenas uma pequena fração dos bits, e introduzir erros menores, ou ainda fazer clones imperfeitos dos qubits transmitidos por Alice, de forma a possuir alguma correlação entre sua seqüência de bits e a seqüência compartilhada por Alice e Bob. Para evitar que

Eva ainda possua informação a respeito da chave final, alguns algoritmos clássicos de *correção de erros* e *amplificação de privacidade* são utilizados[28].

Podemos, portanto, resumir o protocolo BB84 da seguinte forma:

1. Alice escolhe duas seqüências de bits, $\{a_k\}$ e $\{b_k\}$, geradas de forma aleatória, cada uma contendo a mesma quantidade *n* de bits.

2. Para cada *k* variando de 1 a *n*, Alice prossegue da seguinte forma: se $a_k = 0$, então a base $(|0\rangle,|1\rangle)$ é escolhida; se $a_k = 1$, a base $(|+\rangle,|-\rangle)$ é escolhida. Após realizada a escolha de base, Alice prepara e envia o qubit associado ao $(b_k + 1)$-ésimo elemento da base pelo canal quântico. Por exemplo, no caso da base escolhida ter sido $(|+\rangle,|-\rangle)$, $b_k = 1$ corresponde ao envio do qubit $|-\rangle$.

3. Bob recebe os *n* qubits enviados. Para *k* variando de 1 a *n*, Bob escolhe um número aleatório c_k. Se $c_k = 0$, ele realiza uma medida no bit $|k\rangle$ utilizando a base $(|0\rangle,|1\rangle)$, mas se $c_k = 1$ a medida é realizada na base $(|+\rangle,|-\rangle)$. O resultado das medidas é armazenado em uma seqüência de bits $\{d_k\}$.

4. Alice e Bob anunciam, via comunicação clássica, suas seqüências de escolha de base, $\{a_k\}$ e $\{c_k\}$, e computam a seqüência $\{s_k\} = \{a_k\} \oplus \{c_k\}$, onde $\oplus$ é a soma módulo 2. Para cada *k* satisfazendo $s_k = 0$, Alice e Bob guardam os valores b_k e d_k em novas seqüências de bits $\{e_j^{A,B}\}$, que muito provavelmente conterão ~ *n*/2 elementos. Para os elementos onde $s_k = 1$, os termos

[28] Uma idéia que funcionaria, mas que na prática é muito ineficiente, para amplificação de privacidade consiste em selecionar pares de bits ao acaso na seqüência de cada um e calcular sua soma módulo 2 (XOR). É possível mostrar (ver [16] para uma discussão qualitativa) que, realizando-se seqüencialmente essa operação, a informação de Eva diminui progressivamente.

correspondentes b_k e d_k são descartados. A seqüência $\{e_j\}$ é chamada de *sifted key*, e o processo para obtê-la é chamado *sifting*[29].

5. Para verificar a presença de Eva, Alice seleciona, aleatoriamente, um subconjunto da seqüência $\{e_j^A\}$ e anuncia os *valores* (j, e_j^A) para Bob, que os compara com seus valores (j, e_j^B). Se houver muitos erros (ou seja, muitos valores de *j* para os quais as seqüências diferem), o protocolo é interrompido, significando que possivelmente há um espião; caso contrário, Alice e Bob concordam que $\{e_j^A\} \cong \{e_j^B\}$ e guardam todos os valores de bits que não foram anunciados publicamente.

6. Em todos os bits restantes, Alice e Bob aplicam protocolos (clássicos) de correção de erro e amplificação de privacidade, de forma que a informação de Eva é reduzida tanto quanto se queira. No final, Alice e Bob compartilham seqüências idênticas contendo $m < n/2$ bits. Essa chave é chamada de *chave secreta compartilhada*.

Não podemos nos esquecer, entretanto, que o canal quântico pode introduzir ruído e que nem sempre a presença de erros é devida a um espião. Além disso, uma série de imperfeições na implementação prática podem ser fontes de erro. Isso significa que, na prática, uma presença excessiva de erros pode comprometer o sucesso dos algoritmos de correção de erros e amplificação de privacidade. A seção a seguir discute esses problemas.

2.5.4. Implementação prática

Em 1992, os primeiros resultados experimentais de criptografia quântica foram publicados [24]. Nesse primeiro experimento, os qubits foram codificados na polarização de fótons gerados por um LED atenuado e a troca de chaves foi realizada ao longo de uma distância de 30 cm dentro do laboratório.

[29] Embora tratem-se de duas seqüências distintas, uma pertencendo a Alice e a outra a Bob, podemos imaginar se tratar de uma mesma seqüência compartilhada na qual há "erros".

Nos últimos 15 anos, a tecnologia evoluiu de forma estrondosa, e trocas de chave secreta foram estabelecidas a distâncias de centenas de km via fibras ópticas, ou a dezenas de km via espaço livre. Inclusive já é possível encontrar sistemas de criptografia quântica à venda, como é o caso da empresa IdQuantique[30].

Qualquer implementação prática de um sistema de criptografia quântica deve buscar maximizar duas quantidades, que podemos compreender perfeitamente bem apenas usando o bom senso: a *distância máxima* que Alice e Bob podem estar separados um do outro para que uma chave secreta possa ser gerada e a *taxa máxima* de bits secretos por segundo que pode ser obtida. Afinal, ninguém gostaria de adquirir um sistema no qual chaves secretas só podem ser geradas entre indivíduos a poucos metros de distância ou a uma taxa muito lenta (nesse caso, um carro-forte contendo vários DVDs com chaves secretas seria muito mais apropriado!).

Mas o que limita a distância máxima e a taxa de geração de bits secretos? Os vilões são os mesmos dos sistemas de comunicação clássica, só que em versão quântica: as *perdas* e o *ruído*. Perdas surgem por vários motivos; por exemplo, se um fóton se propaga pela atmosfera, ele pode ser absorvido por alguma molécula, ou espalhado por uma partícula de aerosol, ou simplesmente passar direto pelo detector devido a erros de apontamento. É evidente que, quanto maior a distância, maior a probabilidade de um fóton não chegar a seu destino. O ruído também pode surgir de várias fontes, como por exemplo (ainda no caso da propagação atmosférica) luz do sol espalhada pela atmosfera que atinge o receptor, ou (de forma geral) contagens de escuro do detector. Esses dois fatores, perdas e ruído, ao agirem em conjunto, podem limitar severamente o desempenho do protocolo BB84. Uma forma de medir esses dois efeitos ao mesmo tempo, e de simples interpretação física, é a chamada *taxa de erro de qubit* (QBER, *Quantum Bit Error Rate*), que é de extrema importância a qualquer sistema prático de distribuição quântica de chaves e é discutido a seguir. Em seguida, mostramos qual a relação entre a QBER e a segurança de um sistema de criptografia quântica e de que forma seu desempenho é afetado.

[30] Para maiores informações sobre a empresa, ver www.idquantique.com

2.5.4.1.
Taxa de erro de qubit (QBER)

A QBER é definida como a razão entre a probabilidade de se obter uma contagem falsa e a probabilidade total de haver contagens, medida por pulso. Apenas por razões pedagógicas, consideremos, sem perda de generalidade[31], um esquema de detecção como o da figura 2(b). Nesse caso, temos que:

$$QBER \equiv \frac{p_{false}}{p_{total}} = \frac{p_{opt} p_{phot} + p_{noise}}{p_{phot} + 2 p_{noise}} \tag{2.43}$$

Nesta definição, p_{phot} é a probabilidade de se detectar um fóton. Essa probabilidade pode ser escrita da forma $p_{phot} = \mu \eta_{link} \eta_{\det}$, onde cada termo representa, respectivamente, o número médio de fótons por pulso, a transmissão do enlace e a eficiência do detector.

As probabilidades p_{opt} e p_{noise} representam as fontes de contagens falsas. A primeira consiste na probabilidade de um fóton ir parar no detector errado, devido a um contraste imperfeito de polarização (no caso de codificação em polarização) ou a uma visibilidade interferométrica abaixo de 100% (no caso de codificação em *time bins*); a segunda é a probabilidade de ocorrência de contagens de escuro e/ou contagens devido a fótons provenientes de fontes de ruído externas.

Observe que o termo p_{noise} surge no denominador multiplicado por um fator 2 devido ao fato de haver dois detectores contribuindo para a existência de contagens. No entanto, no numerador, o fator 2 desaparece pois o ruído só produz erros na metade do tempo, isto é, apenas nos casos em que ocorre no detector "errado". Claramente, a QBER corresponde à fração de bits errados na chave compartilhada por Alice e Bob após o processo de *sifting*.

Nos casos em que $p_{noise} << p_{phot}$, podemos escrever (2.43) como:

$$QBER \cong p_{opt} + \frac{p_{noise}}{\mu \eta_{link} \eta_{\det}} \equiv QBER_{opt} + QBER_{\det} \tag{2.44}$$

[31] A generalidade aqui se refere ao tipo de implementação (polarização, time bins, etc).

Ou seja, a QBER pode ser aproximadamente escrita como uma soma de uma componente "óptica" e uma componente de "detecção". Observe que (2.44) só é válida se o valor da QBER for de apenas alguns porcento; isso é sempre verdade para a componente óptica, que raramente ultrapassa 1%, tanto em esquemas de polarização como de *time-bins*, mas nem sempre pode ser assumido para a componente de detecção. Nesses casos, nos quais o ruído de detecção é predominante, a QBER óptica pode ser desprezada, como veremos no capítulo 6.

2.5.4.2. Critérios de segurança

Vimos na seção 2.5.3 que o protocolo BB84 é perfeitamente seguro na ausência total de ruído, mas ao mesmo tempo sabemos que esse cenário está muito longe da realidade. Agora que introduzimos uma forma de se medir o ruído (QBER), podemos discutir uma condição suficiente para a segurança da criptografia quântica em sistemas reais.

A idéia básica por trás de qualquer discussão sobre segurança está no fato que Alice, Bob e Eva, em algum momento, fazem medidas em seus qubits. Após as medidas, cada um deles possui um conjunto de variáveis aleatórias, que chamaremos respectivamente de α, β e ε. Podemos, assim, definir uma distribuição de probabilidade conjunta $P(\alpha,\beta,\varepsilon)$ e nos fazer a seguinte pergunta: quais condições $P(\alpha,\beta,\varepsilon)$ deve satisfazer de forma que Alice e Bob possam extrair uma chave secreta?

Esse é um problema matematicamente desafiador e sua solução é completamente não-trivial. No entanto, se aceitarmos como corretas as afirmativas abaixo, que na verdade são teoremas que podem ser demonstrados, é possível solucionar o problema de forma extremamente simples. São elas:

(1) Para uma dada distribuição de probabilidade $P(\alpha,\beta,\varepsilon)$, é possível estabelecer uma chave secreta, usando apenas correção de erros e amplificação de privacidade clássicos, se e somente se $I(\alpha,\beta)\geq I(\alpha,\varepsilon)$ ou $I(\alpha,\beta)\geq I(\beta,\varepsilon)$, onde $I(\alpha,\beta)$ é a informação mútua média entre α e β por qubit.

(2) $I(\alpha,\beta)+I(\alpha,\varepsilon)\leq 1$, isto é, Eva e Bob não podem receber mais informação do que foi enviado por Alice.

Observe que ambas as afirmações (1) e (2) são altamente razoáveis, mesmo que não nos demos ao trabalho de prová-las. Usando esses fatos, podemos deduzir que:

$$I(\alpha,\beta)\leq \frac{1}{2} \tag{2.45}$$

Além disso, podemos escrever $I(\alpha,\beta)=1-h(QBER)$, onde h é a função entropia binária. Juntamente com (2.45), obtemos:

$$\begin{aligned} &QBER\log QBER+(1-QBER)\log(1-QBER)\leq \frac{1}{2} \\ &\Rightarrow QBER\leq 11\% \end{aligned} \tag{2.46}$$

Esse resultado é conhecido como o limite para "segurança incondicional". Na realidade, foi recentemente demonstrado [25] que esse limite pode ser ligeiramente aumentado para $QBER\approx 12\%$ se um certo pré-processamento for realizado antes da aplicação dos algoritmos de correção de erro e amplificação de privacidade. Usaremos esse valor de 12% na seção a seguir.

2.5.4.3. Máxima distância e taxa de geração de chave

As duas principais figuras de mérito de um sistema de criptografia quântica podem ser calculadas a partir da QBER. Vamos inicialmente considerar a taxa de geração de chave, que também chamamos de *taxa líquida*. Ela é dada por:

$$R_{net}=R_{sift}\,F\left[I(\alpha,\beta);I(\alpha,\varepsilon)\right] \tag{2.47}$$

A função F depende do algoritmo utilizado para correção de erros e amplificação de privacidade, que intuitivamente supomos ser uma função das

informações mútuas relevantes. R_{sift} representa a taxa de geração de chave após o processo de *sifting*,

$$R_{sift} = \frac{1}{2} f\, p_{phot} = \frac{1}{2} f \mu \eta_{link} \eta_{\det} \tag{2.48}$$

onde f é a taxa de repetição do laser utilizado por Alice, em pulsos por segundo. A combinação de (2.47) e (2.48) torna evidente a dependência da taxa de geração de chave com relação às perdas no canal, à eficiência na detecção e a quão rapidamente Alice modula seu laser. A dependência com a QBER se torna evidente quando escrevemos $I(\alpha,\beta) = 1 - h(QBER)$, da mesma forma que na seção anterior, onde h é a função entropia binária.

Observe que o termo $F\left[I(\alpha,\beta); I(\alpha,\varepsilon)\right]$ depende não apenas dos algoritmos utilizados como também da estratégia de espionagem utilizada por Eva, e que, portanto, não é possível fornecer uma expressão geral para a taxa de geração de chave.

A *distância máxima* corresponde às máximas perdas no canal quântico que o sistema é capaz de suportar. É evidente de (2.44) que a QBER e a transparência do canal estão intimamente relacionadas; além disso, conforme vimos anteriormente, não é possível estabelecer uma chave secreta se a QBER estiver acima de ~ 12%, de forma que a distância do enlace está limitada por esse valor. Observando que não podemos utilizar a aproximação (2.44) nesse caso mas, ao mesmo tempo, desprezando a QBER óptica, temos de (2.43) que:

$$QBER \cong \frac{1}{2 + \dfrac{\mu \eta_{link} \eta_{\det}}{p_{noise}}} < 12\% \tag{2.49}$$

Rearrumando os termos, obtemos:

$$\eta_{link} > \frac{6.3\, p_{noise}}{\mu \eta_{\det}} \tag{2.50}$$

Considerando que $\eta_{link} = e^{-\alpha L}$, onde α é um coeficiente de atenuação e L a distância entre Alice e Bob, obtemos finalmente:

$$L < \frac{1}{\alpha} \ln\left(\frac{6.3 p_{noise}}{\mu \eta_{\det}}\right) \equiv L_{\max} \tag{2.51}$$

A expressão (2.51) nos fornece todos os parâmetros para maximização da distância máxima. Em alguns deles não temos nenhum controle, como por exemplo o coeficiente de atenuação α; no entanto, podemos facilmente alterar o valor médio de fótons por pulso μ se o desejarmos. Observe que, se fizermos em (2.50) um valor de μ razoavelmente elevado, a transmissão do enlace pode ser tão pequena quanto se queira!

Infelizmente não podemos aumentar μ, muito pelo contrário; o tópico a seguir explica o porquê.

2.5.4.4. Pulsos multi-fóton e o ataque PNS

O que acontece se o número médio de fótons por pulso for aumentado? De acordo com (2.26), a probabilidade de pulsos multi-fóton também aumenta. Não deve ser difícil, a essa altura, perceber por que não podemos pagar esse preço.

Inicialmente, suponha que Eva seja capaz de medir o número de fótons presentes em cada pulso sem introduzir nenhum distúrbio nos qubits (e sem destruir nenhum fóton). Em seguida, a presença de mais de um fóton no mesmo pulso for detectada, Eva guarda um deles para si e permite que o restante prossiga na direção de Bob, caso contrário ela bloqueia o pulso por completo. Dado que em um pulso multi-fóton produzido por um laser todos os fótons codificam a mesma informação, essa situação é equivalente a produzir clones perfeitos! Esse ataque, chamado de *photon number splitting* (PNS), não introduz nenhuma espécie de distúrbio nos qubits recebidos por Bob, ao contrário do que ocorre no ataque de interceptação-reenvio. Desde que possua meios para isso, Eva pode guardar seus fótons em uma "memória quântica" (ver discussão adiante) e postergar sua medida para o momento em que Alice anunciar quais bases usou para codificar seus qubits, de forma a obter informação sobre 100% da chave!

Observe que o ataque ainda não é perfeito pois, a princípio, Alice e Bob poderiam detectar a presença de Eva devido à atenuação provocada pelo ataque PNS. Para isso, bastaria que Eva utilizasse um outro canal, com menor atenuação, para transmitir os qubits dos pulsos multi-fóton a Bob, de forma que as perdas no ataque sejam compensadas. Apesar da viabilidade a tal ataque ser duvidosa[32], ela mostra claramente que a probabilidade de Alice gerar pulsos com 2 ou mais fótons deve ser diminuída ao máximo.

Uma possível forma de aumentar a robustez de um sistema de criptografia quântica contra ataques PNS é alterar o protocolo BB84 para este fim, como por exemplo o protocolo SARG [26]. O setup experimental é idêntico, sendo a única diferença o procedimento de sifting.

Suponha que Alice prepara um dos quatro estados usuais do BB84, por exemplo o estado $|0\rangle$. Agora, ela anuncia a Bob um conjunto de dois estados não-ortogonais que contenha o estado que foi enviado (no mesmo exemplo, ela poderia anunciar o conjunto $\{|0\rangle,|+\rangle\}$). Sempre que Bob medir na base $\{|0\rangle,|1\rangle\}$ (a base correta), encontrará $|0\rangle$ e não poderá concluir nada. Porém, quando fizer uma medida na base $\{|+\rangle,|-\rangle\}$, que é a base *incorreta*, na metade dos casos ele encontrará $|-\rangle$, que ele sabe que não foi enviado por Alice. Logo, nesses casos, ele infere que Alice enviou $|0\rangle$. O ataque PNS é muito menos eficiente no protocolo SARG pois Eva não tem como distinguir entre 2 estados não-ortogonais de forma confiável.

Observe, no entanto, que são perdidos o dobro de bits no processo de sifting, de forma que a taxa de geração de chave secreta cai pela metade com relação ao protocolo BB84.

Uma outra idéia desenvolvida para atacar o mesmo problema foi proposta por Hwang [27], através dos chamados *decoy states* ("estados isca"). Nesse esquema, Alice intencionalmente produz pulsos multi-fóton e os mistura aleatoriamente aos pulsos enviados a Bob. Após a transmissão, durante o

[32] Em primeiro lugar, o ataque depende de uma medida *quantum non-demolition*, muito difícil de um ponto de vista técnico; em seguida, o canal quântico de menor perda não pode ser trivialmente implementado.

processo de reconcicliação de base, eles verificam se houve perdas anormais para esse tipo de pulso. Se as perdas nos estados "isca" forem maiores que nos demais estados "normais", o protocolo é interrompido.

2.5.4.5. Cuidados especiais na realização experimental

Quais são, portanto, as considerações práticas que devem ser levadas em conta no projeto de um sistema de criptografia quântica? Além de todos os aspectos indicados nas seções 2.5.1 a 2.5.3, que são aplicáveis a qualquer sistema de comunicações quânticas, a criptografia quântica exige alguns cuidados extras.

O *transmissor* deve ser capaz de ser modulado rapidamente, de forma que a taxa de repetição *f* introduzida em (2.48) seja a maior possível e, conseqüentemente, o mesmo ocorra com a taxa de geração de chave secreta R_{net}. Observe que essa restrição se impõe muito mais ao elemento responsável pela escolha do estado no qual o qubit será enviado do que ao laser que gera os pulsos atenuados. No caso de codificação em polarização, uma forma de evitar a utilização de componentes ativos consiste na utilização de 4 lasers, conforme sugerido na figura 51 do capítulo 5. Neste caso, convém realizar uma filtragem espectral, de forma que não seja possível reconhecer qual laser produziu um dado fóton.

O *receptor* também deve ser capaz de rapidamente alterar sua base de medida, ou então utilizar uma escolha passiva de base usando quatro detectores – analogamente ao uso de 4 lasers na transmissão – de forma semelhante ao ilustrado na figura 52 do capítulo 5. Além disso, de forma a minimizar a QBER de detecção, a probabilidade de ruído deve ser a menor possível; para isso, é imprescindível que a duração dos pulsos seja pequena (~ 1 ns) e que o ruído de detecção não seja muito elevado.